SCIENCE AND TECHNOLOGY RESEARCH CENTRE
UNIVERSITAS PGRI SEMARANG - INDONESIA

ISSN: 2715-4211

Journal ASSET

Advance Sustainable Science, Engineering and Technology

Volume 6, Number 2 (2024): April

URL LINK: http://journal.upgris.ac.id/index.php/asset

Advance Sustainable Science, Engineering and Technology

e-ISSN: 2715-4211

Volume 6 Number 2 April 2024

Science and Technology Research Centre, Universitas PGRI Semarang, Indonesia

Advance Sustainable Science, Engineering and Technology (ASSET) is a peer-reviewed open-access international scientific journal dedicated to the latest advancements in sciences, applied sciences and engineering, as well as relating sustainable technology. This journal aims to provide a platform for scientists and academicians all over the world to promote, share, and discuss various new issues and developments in different areas of sciences, engineering, and technology.

ASSET papers will be published twice a year by Universitas PGRI Semarang, Central Java, Indonesia.

Advance Sustainable Science, Engineering and Technology
e-ISSN: 2715-4211
Volume 6 Number 2 April 2024
Science and Technology Research Centre, Universitas PGRI Semarang, Indonesia

Editorial Preface

Advance Sustainable Science, Engineering and Technology (ASSET)
Volume 6 Number 2 April 2024

We are delighted to present Volume 6 Number 2 of Advance Sustainable Science, Environmental Engineering, and Technology (ASSET). This issue comprises 25 manuscripts that showcase a diverse range of research studies across various fields.

The publication features insightful research, including a spatial analysis of accessibility levels for the Bobo Village community in the Lore Lindu National Park area, the optimization of biogas liquid waste from livestock manure for renewable energy through Microbial Fuel Cell (MFC) technology, and a comparison of optimizer strategies to enhance emotion classification in IndoBERT models.

Additionally, this volume includes studies on implementing a data layer in a blockchain network using the SHA256 hashing algorithm, aspect-based sentiment analysis on electric motorcycles, and the identification of chatbot usage in online store services using natural language processing methods. The publication also covers the implementation of the Adaboost method to increase the accuracy of early diabetes predictions based on decision trees, as well as the identification of landslide-prone areas using the Schlumberger configuration geoelectric method.

Other topics addressed in this volume include enhancing web server security against layered cyber threats in healthcare, implementing blockchain for publishing and verifying digital certificates in EduTech, and a bottled water distribution system using web-based distribution requirement planning and trend moment algorithms. Supplier selection modeling and analysis in the metal casting industry using Analytical Hierarchy Process, prediction of the birth rate of babies at regional hospitals using the Naïve Bayes algorithm, and the development of website-based stunting prevention educational media services are also covered.

Furthermore, this volume includes studies on utilizing data mining techniques to analyze changes in the purchase behavior of batik customers, analyzing hero performance in Mobile Legends with the K-Nearest Neighbors approach, and project delay factor analysis using the Fault Tree Analysis (FTA) method and rescheduling with the Critical Path Method (CPM) in ducting manufacturing projects. Additionally, it discusses protein concentrate from tuna head waste, MERN implementation in online quiz applications to recognize and avoid social media hoaxes, and analysis of improving the quality of Fatty Acid Methyl Ester (FAME) products against Acid Value (AV) levels using the Six Sigma and Kaizen methods.

The publication also covers the optimization of sample warehouse layout, habitat suitability analysis for Macaca Tonkeana, sustainability strategies of traditional Vannamei shrimp cultivation, design of a mobile-based wedding information and booking system using Backend as a Services (BaaS) on the Android platform, and a portable low-cost home sleep monitor using Wemos D1 Mini.

We extend our sincere thanks to all 83 authors from various esteemed institutions including Tadulako University, Tadulako, Central Sulawesi; Institut Sains Teknologi dan Kesehatan 'Aisyiyah Kendari, Southeast Sulawesi; Universitas Halu Oleo, Kendari, Southeast Sulawesi; Universitas Siliwangi, Tasikmalaya, West Java; Universitas Dian Nuswantoro Semarang, Central Java; Politeknik STMI Jakarta,

Jakarta; Institut Teknologi dan Bisnis Nobel Indonesia; Universitas Pamulang, Banten; Universitas Islam Negeri Sunan Kalijaga, Yogyakarta; Universitas Muhammadiyah Gresik, Gresik; Universitas Pembangunan Nasional "Veteran" Jawa Timur, Surabaya; Universitas Kristen Krida Wacana, Jakarta; Satya Wacana Christian University; BMKG (Badan Meteorologi, Klimatologi, dan Geofisika), Indonesia; and Indian Institute of Technology Jodhpur, Jodhpur, India.

We hope that this publication will serve as a valuable resource for researchers, academicians, and practitioners in advancing sustainable science, engineering, and technology. Through these collective efforts, we aim to contribute significantly to the body of knowledge and promote sustainable practices in various fields. We look forward to continued collaboration and the exchange of ideas in future editions of ASSET.

Thank you once again to all the authors, reviewers, and contributors for their valuable contributions to this volume.

April 2024
Assoc. Prof. Mega Novita
Asst. Prof. Rizky Muliani Dwi Ujianti

Advance Sustainable Science, Engineering and Technology
e-ISSN: 2715-4211
Volume 6 Number 2 April 2024
Science and Technology Research Centre, Universitas PGRI Semarang, Indonesia

Table of Content

Advance Sustainable Science, Environmental Engineering and Technology (ASSET)
Volume 6 Number 2 April 2024

Advance Sustainable Science, Engineering and Technology
e-ISSN: 2715-4211
Volume 6 Number 2 April 2024
Science and Technology Research Centre, Universitas PGRI Semarang, Indonesia

Advance Sustainable Science, Engineering and Technology (ASSET)
Vol. 6, No.2, April 2024, pp. 0240201-01 ~ 0240201-09
ISSN: 2715-4211 DOI: https://doi.org/10.26877/asset.v6i2.18000

Spatial Analysis of the Level of Accessibility of the Bobo Village Community in the Lore Lindu National park Area

Nurul Istiqamah*, Hasriani Muis, Sudirman Daeng Massiri, Rhmadhani Fitrah Baharuddin, Misrah, Bau Toknok, Arman Maiwa, Abdul Rahman, Amati Eltriman Hulu

Faculty of Forestry, Tadulako University, Tadulako, Jl. Soekarno Hatta Km. 9 Palu 94119, Central Sulawesi, Indonesia

*nurulistiqamahabu@gmail.com

Abstract. This research aims to determine the level of community accessibility in the Lore Lindu National Park Conservation area and determine the level of community accessibility in reaching areas that have been selected based on the provisions of the Cooperation Agreement (PKS) between Lore Lindu National Park and the Bobo Village Government in 2021. This research uses spatial analysis with five variables: land use, slope, and distance from roads, rivers and settlements. Determining the weight of each variable uses the Analytical Hierarchy Process (AHP) method with sources from the Bobo Village community and the Bobo Village Conservation Management Institute. The research results show that with the AHP method, land use has a significant effect on the level of community accessibility with a weight of 52.3%, followed by other parameters, namely slope of 26.8%, distance from the main road of 11.3%, distance from the river is 3, 4%, and distance from settlements is 6.2%. Furthermore, three classes of accessibility levels for the Bobo Village community were obtained: low level covering an area of 390.94 ha (96%), medium level covering an area of 17.76 ha (4%) and high level surrounding an area of 0.11 ha. (0%).

Keywords: Accessibility, Community, Conservation area, Analytical Hierarchy Process

(Received 2024-01-01, Accepted 2024-01-12, Available Online by 2024-02-08)

1. Introduction

As one of the largest archipelagic countries in the world, Indonesia has extraordinary natural wealth [1], including forest areas that play an important role in preserving the environment and human life. People's dependence on forest areas in Indonesia is reflected in various aspects of daily life. Economically, forests provide abundant natural resources, such as wood, non-timber forest products and traditional medicinal plants [2]. Local communities often depend for their livelihoods on activities such as agriculture, plantations and grazing around forest areas [3].

Community access to forests is important in maintaining the balanced use of natural resources. According to [4], communities around forests often have rights and traditions in utilizing forest

resources to meet their daily needs. However, it should be noted that forest use that is not well managed can contribute to higher deforestation rates [5]. Legal access to the community refers to recognizing access rights to natural resources as regulated by applicable laws or regulations [6]. This step is carried out to regulate and facilitate the community's use of natural resources per established regulations.

The level of accessibility of forest communities is an important factor that enables communities to efficiently manage and utilize the natural resources available in forest areas [7]. Communities around forests often face challenges with accessibility to forest resources, which can affect their welfare [8]. Several findings show that the lack of community access to forests, the large amount of critical land that has not been utilized in and around forest areas [9], and the lack of incentives provided by the government to forest communities can cause poverty in villages around forests to continue [10].

Based on the Decree of the Minister of Forestry Number SK.869/Menhut-II/2014, Lore Lindu National Park (TNLL) is located in Central Sulawesi and has an area of around 215,733.7 Ha [11]. TNLL applies the concept of community-based collaborative management as a reference in managing conservation areas as outlined in the Decree of the Director General of KSDAE Number 456/KSDAE/SET.REN.2/8/2017. In 2018, through a conservation partnership program, the Lore Lindu National Park Center provided the community access to non-timber forest resources and environmental services in the conservation area [12]. In 2021, the TNLL Center, with the support of Forest Program III Sulawesi, became a facilitator by taking a conservation partnership approach so that 56 Conservation Partnership Cooperation Agreements (PKS) were established between the village government and the Lore Lindu National Park Center [13].

Bobo Village is a village in Palolo District, Sigi Regency, Central Sulawesi Province and is one of the supporting villages of Lore Lindu National Park [13]. Bobo Village has collaborated with the TNLL Center, which regulates conservation partnerships in the context of community empowerment through providing access to the collection of Non-Timber Forest Products, namely sugar palm and candlenuts and Traditional Cultivation in the traditional zone of Lore Lindu National Park [14]. This aims to increase community income without disturbing the sustainability of conservation areas [15].

Communities around conservation areas need to be considered because the community still depends on conservation areas. [12] stated that the use of forest resources by local communities in the Lore Lindu National Park conservation area is still relatively high even though formal regulations prohibit the exploitative use of forest resources. The presence of the community is a threat factor in the management of conservation areas, such as encroachment, shifting cultivation, gold mining without permits (PETI) and so on [16].

This research on spatial analysis of the level of community accessibility is important to determine the level of community accessibility in the Lore Lindu National Park Conservation area and to assess the extent of community accessibility in reaching areas that have been determined by the provisions of the Cooperation Agreement (PKS) using the Geographic Information System (SIG) [17]. This research can evaluate whether the community accesses the Community Conservation Agreement (KKM) area by the agreement or whether there is a tendency to carry out activities outside the area boundaries stipulated in the agreement.

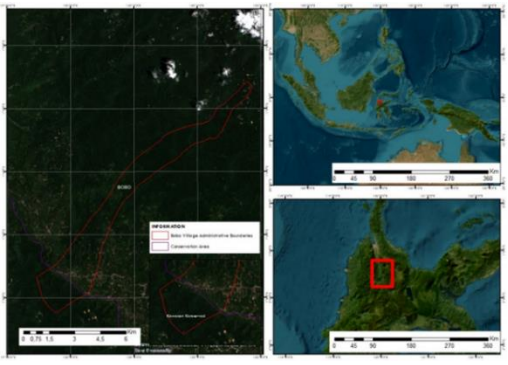


Figure 1. Map of research location

2. Research methods

2.1 Research Location and Time

The research location covers the administrative area of Bobo Village, which is geographically located at 120°00'11.0" East Longitude and 1°07'18.1" South Latitude. The area of Bobo Village is 408.83 ha. The research was conducted from August to December; the stages began with distributing questionnaires to respondents, processing data, and writing up research results.

2.2 Tools and Materials

The tools used in this research were writing tools, calculating tools, and a set of computers, ArcGIS version 10.8, Global Positioning System (GPS), camera, and Microsoft Office. The material used is Sentinel imagery. Other data used are a digital map of Bobo Village's landform, an administrative map of Bobo Village, a road access map, and a slope map.

2.3 Data Collection and Analysis

Data collection in this research is by field measurements/observations and interpretation of maps. The types of data are land use data sourced from Citra Sentinel 2A, slope sourced from DEMNAS, distance from roads sourced from the Indonesian Earth Map (RBI), distance from rivers sourced from the Indonesian Earth Map (RBI), and distance from settlements sourced from the Indonesian Earth Map (RBI).

Data analysis includes analysis of variables assessing the level of accessibility of the Bobo Village community in the Lore Lindu National Park Area, spatial analysis, and attribute analysis. Analysis of variables evaluate the level of accessibility of the Bobo Village community in the Lore Lindu National Park area was carried out by analyzing Sentinel 2A imagery and buffer analysis of the Indonesian Earth Map (RBI) map in the Lore Lindu National Park area, which consisted of distance from the road, distance from the river and distance from the settlement. Spatial analysis and attribute analysis include classifying each variable, giving a score, then overlaying a map of the analysis results of each variable using the specified formula. The final result of the analysis is a map of the accessibility level of the Bobo Village community in the Lore Lindu National Park Area based on the formula used.

2.4 Analysis of Accessibility Level Variables

2.4.1. Land Use Analysis

Land Use Analysis is an interaction between humans and the environment, which is characterized by human efforts to utilize land for their lives and needs [18]. Land use classification can determine the type of land use in an area and assess the level of community accessibility through activities carried out in a location based on area Community Arable land. Land use classification seen from Sentinel 2A Image, which was downloaded on the website https://scihub.copernicus.eu/ on July 8, 2023, then processed using ArcGIS 10.8 with composite band 432 (natural color). The classification results are cut based on the research area.The community accessibility level classes from the scoring can be seen in Table 1.

2.4.2. Slope Analysis

The specific height difference on the ground surface that forms an area is called the slope [4]. Research [19] shows that the slope can influence people's accessibility to conservation areas; the flatter or sloping an area is, the more dominant people are in carrying out activities in that area. The class levels of community accessibility to the slope can be seen in Table 1.

2.4.3. Distance Analysis from the Road

One factor that influences the level of accessibility is road infrastructure. The existence of a road network around or within the area allows people to access and participate in various activities within the region. [20] research results show that the proximity of forest areas to the road network

influences community accessibility. The following classes of community accessibility levels regarding distance from the road can be seen in Table 1.

2.4.4. Distance Analysis from the River

Rivers are one of the factors for community accessibility in national park areas. The existence of a river system around or within the area can open access for the community to engage in various activities, such as agriculture, plantations, and others [21]. The following classes of community accessibility levels regarding distance from the road can be seen in Table 1.

2.4.5. Distance Analysis from Settlements

The existence of settlements around national parks can affect people's accessibility to the area. Settlements act as the main access or starting point for people to enter the national park. The existence of settlements close to the area encourages and makes it easier for people to carry out activities [22].

Table 1. Community Accessibility Level variable assessment

Land Use		**Slope**		**Distance from Road**	
Class	**Scoring**	**Class %**	**Scoring**	**Class (Meters)**	**Scoring**
Settlement	5	0-8	5	0-600	5
Dryland farming	4	8-15	4	600-1200	4
Plantation	3	15-25	3	1200-1800	3
Bushy bushes	2	25-40	2	1800-2400	2
Primary Forest	1	>40	1	>2400	1
Distance from River				**Distance from Settlements**	
Class (Meters)	**Scoring**			**Class (Meters)**	**Scoring**
0-600	5			0-600	5
600-1200	4			600-1200	4
1200-1800	3			1200-1800	3
1800-2400	2			1800-2400	2
>2400	1			>2400	1

2.4.6 Variable Weight Analysis Using Pairwise Comparison

In this research, determining the weight of each variable will be carried out using the Pairwise Comparison method, taking into account the impact of each variable on community accessibility in Lore Lindu National Park through the views of a number of experts or stakeholders using a questionnaire. Some of the parties who were resource persons in the research were the Bobo village community, who have activities in the Lore Lindu National Park, and members of the Village Conservation Management Institute.

2.5 Weighting of Community Accessibility Level Variables

The process of weighting each parameter uses the AHP (Analytical Hierarchy Process) method, which involves ranking and rationality comparison. This method requires a hierarchical structure or solid feedback network, including criteria for various influences, stakeholders, and decision alternatives to determine the best choice [23]. In the context of this research, weighting is carried out by evaluating the influence of the parameters Land use, Slope Slope, Distance from roads, Distance from rivers, and Distance from settlements. The determination of weights is based on an analysis of the influence between parameters on the Accessibility Level of the Bobo Village Community in the Lore Lindu National Park Area.

The hierarchical consistency evaluation process is carried out through a consistency test, which involves calculating the Consistency Index (CI) and Consistency Ratio (CR). CR functions as a determinant of whether the calculation is consistent or not. There are two classifications of CR values:

- If $CR < 0.10$, it indicates a fairly rational level of consistency in pairwise comparisons.
- If $CR \geq 0.10$, it indicates inconsistency in the assessment, and it is recommended that a repeat

questionnaire survey be carried out.

Next, the Random Index (RI) value is determined based on the matrix order that has been created, and the RI value for a particular matrix order can be seen in Table 3.

Table 2. Random Index Value

N	1	2	3	4	5	6	7	8	9	10
(RI)	0.00	0.00	0.52	0.89	1.11	1.25	1.35	1.40	1.45	1.49

2.6 Analysis of community accessibility levels

To determine the level of community accessibility, the vector data for each Variable is scored by multiplying the score by the weight of each Variable, and the scoring results are added to a new field in the attribute table. Then, the spatial vectorization results are overlay analysis by combining the vector layers of each Variable into one vector layer. The overlay results are classified based on the level of community accessibility by adding up the results by multiplying the scores and weights for each Variable.

$$Accessibility\ Level = Kb.Kp + Sb.Sp + Pb.Pp + Jb.Jp + Lb.Lp \quad (1)$$

K = Slope
L = Land Use
S = Distance from River
b = Weight Value
P = Distance from Settlements
p = Parameter Class Score
J = Distance from Road

The interval value of the level of community accessibility uses the Sturges formula by dividing the largest data value and the smallest data value [24];

$$Interval\ Class = \frac{(Xt - Xr)}{k} \quad (2)$$

Xt = Largest data *Xr* = Smallest data *K* = Number of classes

3. Results and discussion

3.1 Community Accessibility Level Parameters

3.1.1 Slope

The slope of the slopes in the Lore Lindu National Park (TNLL) area dominates in the range of >40% covering an area of 280.47 Ha; 25-40% covering an area of 81.53 ha (20%); 15-25% covering an area of 29.29 ha (7%); 8-15% covering an area of 13.36 ha (3%); and 0-8% covering an area of 4.18 ha (1%). The steep condition of the area causes low accessibility because it creates topographic conditions that are very steep and difficult to access. Physically [25] states that the steeper the slope, the lower the level of accessibility. Significant slopes can hinder vehicle and pedestrian movement and limit community access to natural resources and local economic activities.

3.1.2 Distance to Road

With adequate road access, the level of community accessibility will be higher. [13] stated that the closer the TNLL area is to the road, the easier it will be for people to access it. Based on the analysis results, it was found that the distance between 0-600 meters covering an area of 195.44 ha (48%), 600-1200 meters covering an area of 132.37 ha (32%), 1200-1800 meters covering an area of 71.06 ha (17%), 1800-2400 meters covering an area of 9.96 ha (2%).

The distance from the road to the TNLL area between 0-600 meters has a percentage of 48% or an area of 195.44 Ha. Even though the distance from the road is relatively short, field facts show that a footpath is the only road access available. This indicates that even though they are geographically close, people still experience difficulties accessing these locations due to the lack of adequate and accessible road infrastructure.

3.1.3 Distance to River

The closer an area is to a river or water source, the higher the possibility of access [26]. However, based on survey results from the Bobo Village community, it was found that the closer they were to the river, the more people did not carry out activities along the river. This is due to the fact that the source of clean water for the people of Bobo Village comes from a conservation area. The Conservation Area has many river networks, so the analysis results show that the entire Conservation Area is at a distance of less than 600 meters.

3.1.4 Distance to Settlement

The distance between settlements and the Lore Lindu National Park area influences the community's access to non-timber forest products. The research results show that the area with a distance of between 600 to 1200 meters from the settlement is a fairly large area with an area of 159.76 ha (39%), and the distance from 0 to 600 meters is 111.67 (27%), 1200-1800 meters each covering an area of 114.39 ha (28%), 1800-2400 meters covering an area of 23.02 ha (6%). Distance from settlements influences the extent to which community access is adequate. This is by the statement [27] that the closer an area is to a settlement, the easier and wider public access will be.

3.1.5 Land Use

The analysis results show that the Lore Lindu National Park area is dominated by primary forest land use of 92% with an area of 376.98 Ha. Significantly, the level of difficulty of accessibility in this area indicates limited potential for human interaction with the forest [28]. Despite this, around 7% of the land is used for dryland agriculture, while settlements only account for 1% of total land use. These results reflect the complex dynamics between forest conservation, agricultural needs, and settlements, which are important to consider in planning natural resource management and ecosystem sustainability.

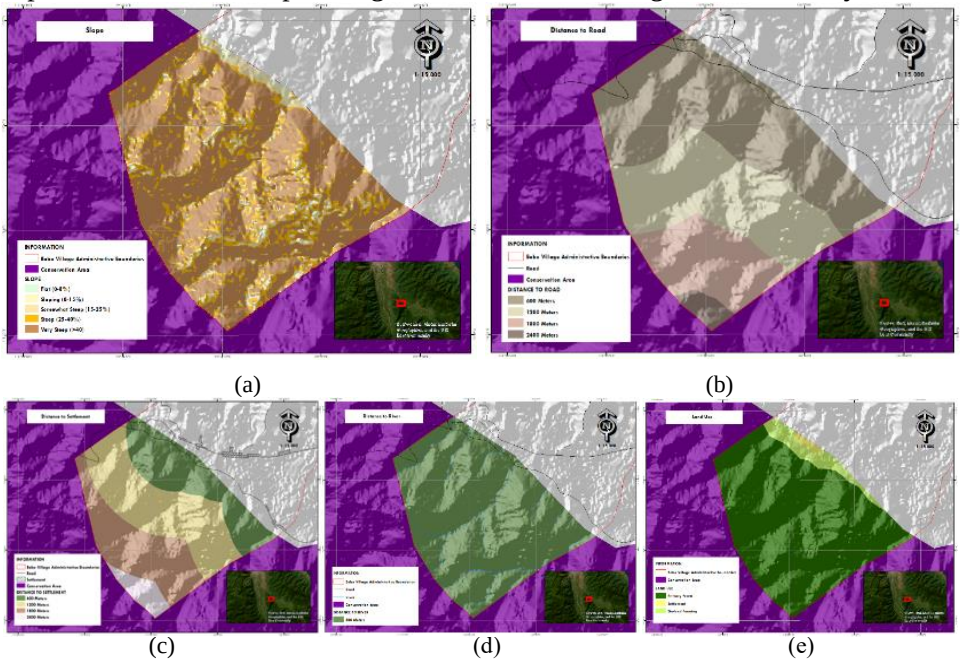


Figure 2. Map of community accessibility level parameters: slope slope (a), distance from roads (b), distance from rivers (c), distance from settlements (d) land use (e)

3.2 Parameter Weighting

Parameter weighting is an analysis in decision-making. This analysis focuses on all parameters, namely the slope map, distribution map of non-timber forest products, residential buffer distance, road buffer distance and river buffer distance, to determine the level of community accessibility in the Lore Lindu National Park Area, which was obtained using the AHP method. The comparison results were

obtained from a questionnaire filled in by seven respondents. The questionnaire data collected was processed using Microsoft Excel to determine the priority weight of each parameter with an inconsistency limit of ≤ 0.1.

The weighting results that will be used are based on the criteria for consistency provisions with the Consistency Ratio (CR) value that must meet the CR requirement ≤ 0.1. Consecutively, the weight of each parameter is land use at 52.3%, slope at 26.8%, distance from the road at 11.3%, distance from settlements at 6.2%, and distance from the river at 3.4%. The results of weighting each parameter of the level of community accessibility that has met these requirements with an inconsistency value of 0.01>0.1.

3.3 Community Accessibility Level

The level of community accessibility in the Lore Lindu National Park Area is analyzed based on scoring and determining the weight of each parameter, namely slope, distance from rivers, distance from settlements, distance from roads and land use, then overlaid, which can be seen in Figure 3.

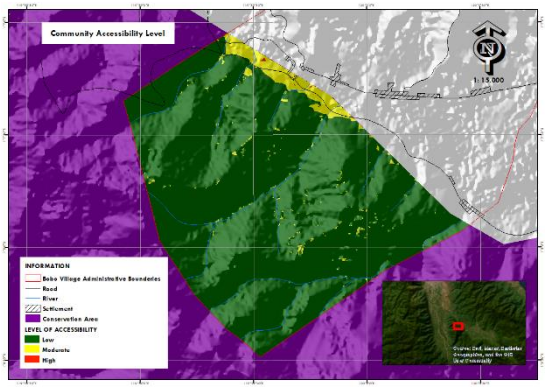


Figure 3. Community Accessibility Level Map

The community accessibility level classes that have been analyzed consist of low, medium and high, which can be seen in Table 3.

Table 3. Level of Community Accessibility in TNLL

No	Accessibility Level	Area (Ha)	Persentase (%)
1.	Low	390.94	96%
2.	Medium	17.77	4%
3.	High	0.11	0%
	Total	408.83	100%

Based on the results of an analysis of the level of accessibility of the Bobo Village community in the Lore Lindu National Park Area, it was found that 96% (390.95 Ha) of the population was in the low category, while only 4% (17.77 Ha) was in the medium category, while the one in the high category had an area of 0.11 Ha (0%). These findings indicate that the majority of residents around the Lore Lindu National Park area face significant obstacles in utilizing resources in accordance with the Cooperation Agreement (PKS) between the Lore Lindu National Park Center and Bobo Village.

The parameters used in this research indicate that the Lore Lindu National Park (TNLL) area is located in an area with very steep slopes and is difficult to access physically. Even though it is geographically close, people still experience difficulties in accessing this location; this is because the only road available is a footpath, so people still have difficulty accessing it. In addition, the distance between settlements and TNLL is relatively close, providing a good opportunity for the community to utilize available resources. However, a striking finding is the existence of rivers within the TNLL area, which results in limited community activities along the river borders because people still depend on rivers as a source of clean water.

Apart from that, the non-timber forest products that can be utilized by the community are very limited, so community activities are not optimal. These findings provide valuable insights for formulating policies and strategies aimed at increasing community accessibility and preserving ecosystems contained in Community Conservation Agreements.

4. Conclusion

The level of accessibility of the Bobo Village community in the Lore Lindu National Park (TNLL) area consists of 3 categories, namely low, medium and high. Of the total research area of 408.83 Ha, 390.95 Ha (96%) is included in the low category, and the level of accessibility is medium at 17.77 Ha (4%). The research results show that the level of community accessibility in the Lore Lindu National Park Area is relatively low. This fact shows the lack of community activity in the Community Conservation Agreement which should be an active area. Therefore, it is necessary to carry out a feasibility study and consultation with the local community in the process of selecting the KKM area so that the community can participate in efforts to utilize the resources controlled and preserve the area.

Acknowledgments

Thank you to the entire academic community of the Faculty of Forestry, Tadulako University, and the Independent Campus Learning Program (MBKM), as well as to the Forestry Student Research Institute (SETMA), which supported this research.

References

[1] M. A. Aminuddin and Agussalim Burhanuddin, "Potensi Kekayaan Dan Keberagaman Maritim Di Wilayah Papua Dalam Upaya Mendorong Kesejahteraan Rakyat," *Mandub : Jurnal Politik, Sosial, Hukum dan Humaniora*, vol. 1, no. 4, pp. 157–176, Nov. 2023.

[2] S. Sulham, S. Umar, A. Maiwa, and A. Rahman, "Identifikasi Manfaat Hutan Oleh Masyarakat Desa Karunia Kecamataan Palolo Kabupaten Sigi," *Jurnal Warta Rimba*, vol. 9, no. 3, pp. 177–182, 2021.

[3] A. C. J. W. Nay, M. E. Pellondo'u, and N. Rammang, "COMMUNITY DEPENDENCE LEVEL ON THE FOREST(Case Study: Forest Area of Koa Besipae, Mio Village, South Amanuban District, South Central Timor Regency)," *Jurnal Wana Lestari*, vol. 5, no. 1, pp. 51–56, 2023.

[4] D. Deandra and A. Tridakusumah, "Akses Masyarakat Dusun Leuwiliang Desa Tanjungwangi terhadap Kawasan Konservasi Taman Buru Masigit Kareumbi Wilayah Kabupaten Bandung," *Jurnal Ekonomi Pertanian dan Agribisnis*, vol. 5, no. 1, pp. 195–203, Jan. 2021.

[5] N. Ramadhany, "Laju Deforestasi Hutan Akibat Aktivitas Pertambangan di Provinsi Kalimantan Timur," *Jurnal Rekayasa Hijau*, vol. 7, no. 1, pp. 10–19, Jul. 2023.

[6] A. T. M. Sihombing and R. Banke, "Politik Hukum Pengelolaan Sumber Daya Alam Di Indonesia," *JUurnal Ilmiah Saintek*, vol. 7, no. 1, pp. 7–15, 2023.

[7] C. Wulandari, S. Bakri, M. Riniarti, and H. J. Putra, "The Level of Encroachers's Acceptance on Community Forest Development Plan: Study at Selagai Linggai Forest Management Unity Regency Central Lampung-Indonesia," *Jurnal Hutan Tropis*, vol. 10, no. 2, pp. 114–123, 2022.

[8] M. M. Haslan and M. Ilyas, "Bentuk Marginalisasi Hak-Hak Masyarakat Lokal Dalam Pengelolaan Sumberdaya Hutan (Kajian Sistem Pengelolaan Sumberdaya Hutan di Pulau Lombok Nusa Tenggara Barat)," *CIVICUS : Pendidikan-Penelitian-Pengabdian Pendidikan Pancasila & Kewarganegaraan*, vol. 10, no. 2, pp. 20–28, 2022.

[9] F. Nabila, P. Satriyo, and I. Ramli, "Identifikasi Distribusi Lahan Kritis Pada Daerah Aliran Sungai (DAS) Krueng Peusangan (Identification of the Distribution of Critical Land in the Krueng Peusangan Watershed)," *Jurnal Ilmiah Mahasiswa Pertanian*, vol. 7, no. 4, pp. 821–831, 2022.

[10] S. M. Utami and A. A. Arsi, "Lingkungan Perdesaan: Sebuah Tantangan Perubahan Bagi Masyarakat Pegunungan," *Bookchapter Alam Universitas Negeri Semarang*, no. 1, pp. 90–120, Aug. 2022.

[11] Menteri Kehutanan, *Kawasan Hutan dan Konservasi Perairan Provinsi Sulawesi Tengah* . 2014.

[12] S. D. Massiri, B. Nugroho, H. Kartodihardjo, and R. Soekmadi, "Institutional sustainability of a community conservation agreement in lore lindu national park," *Forest and Society*, vol. 3, no. 1, pp. 64–76, Apr. 2019, doi: 10.24259/fs.v3i1.5204.
[13] S. Massiri, *Strategi Pemberdayaan Masyarakat pada Kawasan Konservasi*, I. Bandung: Smart Media Publishing, 2022.
[14] "Perjanjian Kerjasama antara Balai Besar Taman Nasional Lore Lindu dengan Kepala Desa Bobo Kecamatan Palolo Kabupaten Sigi Provinsi Sulawesi Tengah." S, pp. 1–8, 2021.
[15] S. daeng Massiri, *Membangun Kesepakatan Konservasi Masyrakat*, Pertama. Palu: Balai Besar Taman Nasional Lore Lindu Kementerian Lingkungan Hidup dan Kehutanan, 2019.
[16] Y. N. Wijayanti, I. U. Tiara, and W. Wijaya, "Faktor Penyebab Perambahan Hutan Di Kaki Gunung Kerinci Dalamkawasan Tnks Di Desa Gunung Labu Kecamatan Kayu Aro Barat," *Jurnal Review Pendidikan dan Pengajaran*, vol. 6, no. 2, pp. 474–482, 2023.
[17] A. E. Hulu *et al.*, "Spatial Analysis of Water Infiltration Potential in the Miu Watershed of Sigi Regency," *Advance Sustainable Science Engineering and Technology*, vol. 5, no. 2, p. 0230208, Jul. 2023, doi: 10.26877/asset.v5i2.16626.
[18] L. Juniyanti, L. B. Prasetyo, D. P. Aprianto, H. Purnomo, and H. Kartodihardjo, "Land-use/land cover change and its causes in Bengkalis Island, Riau Province (from 1990-2019)," *Jurnal Pengelolaan Sumberdaya Alam dan Lingkungan*, vol. 10, no. 3, pp. 419–435, 2020.
[19] H. N. Rahmawati and P. J. Osly, "Identification Of Slope As A Parameter Of Settlement Land Suitability (Case Study Of Bogor," *Jurnal Spektran*, vol. 11, no. 2, p. 163, 2023.
[20] C. Wulandari *et al.*, "The Level of Encroachers's Acceptance on Community Forest Development Plan: Study at Selagai Linggai Forest Management Unity Regency Central Lampung-Indonesia," Cetak, 2022.
[21] M. A. Suni, H. Muis, I. Arianingsih, M. Misrah, and R. F. Baharuddin, "Analisis Dan Pemodelan Spasial Perubahan Tutupan Lahan Di Hutan Produksi Terbatas Kecamatan Kulawi," *Jurnal Tanah dan Sumberdaya Lahan*, vol. 10, no. 2, pp. 273–284, Jul. 2023.
[22] B. Hero Saharjo and S. Khazimastasia, "Early Warning System in Preventing The Forest Fire in KPH Kuningan, West Java," *Jurnal Silvikultur Tropika*, vol. 10, no. 03, pp. 178–183, 2019.
[23] R. W. Saaty, *Decision making in complex environments the Analytic Network Process (ANP) for dependence and feedback. Creative Decisions Foundation, Pittsburgh*. 2016.
[24] H. Rakuasa and P. C. Latue, "Analisis Spasial Daerah Rawan Banjir Di Das Wae Heru, Kota Ambon," *Jurnal Tanah dan Sumberdaya Lahan*, vol. 10, no. 1, pp. 75–82, Jan. 2023.
[25] I. H. Sitorus, F. Bioresita, and N. Hayati, "Analisa Tingkat Rawan Banjir di Daerah Kabupaten Bandung Menggunakan Metode pembobotan dan Scoring," *Jurnal teknik ITS*, vol. Vol.10, No.1, 2021.
[26] I. Madani, S. Bachri, and S. Aldiansyah, "PEMETAAN KERAWANAN BANJIR DI DAERAH ALIRAN SUNGAI (DAS) BENDO KABUPATEN BANYUWANGI BERBASIS SISTEM INFORMASI GEOGRAFIS," *Jurnal Geosaintek*, vol. 8, no. 2, pp. 2502–3659.
[27] P. Fitria, A. Jauhari, and F. Rianawati, "ANALISIS TINGKAT KERAWANAN KEBAKARAN HUTAN DAN LAHAN BERBASIS PENGINDERAAN JAUH DI KECAMATAN KARANG INTAN," *Jurnal Sylva Scienteae*, vol. 4, no. 6, p. 1120, Dec. 2021,.
[28] A. T. Cahyani, A. Program, and S. Geografi, "ANALISIS KESESUAIAN PENGGUNAAN LAHAN TERHADAP ARAHAN FUNGSI KAWASAN KECAMATAN KOTO XI TARUSAN, KABUPATEN PESISIR SELATAN," *JURNAL BUANA*, vol. 4, no. 6, pp. 1178–1188, Nov. 2020, doi: 10.24036/STUDENT.V4I6.1266.

Advance Sustainable Science, Engineering and Technology (ASSET)
Vol. 6, No.2, April 2024, pp. 0240202-01 ~ 0240202-08
ISSN: 2715-4211 DOI: https://doi.org/10.26877/asset.v6i2.17931

Optimization of Biogas Liquid Waste from Livestock Manure as a Source of Renewable Energy through Microbial Fuel Cell (MFC) Technology

Alwahab[1*], Kurnia Sri Yunita[1], Pratitis Tri Maharani[2], Wa Ode Miranda Nur K.[2], La Ode Ahmad[2], Alimin[2], Ahmad Zaeni[2], Ida Usman[2], Prima Endang Susilowati[3]

[1]Faculty of Science Technology and Health, Institut Sains Teknologi dan Kesehatan 'Aisyiyah Kendari, Jl. Kapten Piere Tendean No. 45a Kendari, Southeast Sulawesi, Indonesia

[2]Faculty of Mathematics and Natural Sciences, Universitas Halu Oleo, Jl. H.E.A. Mokodompit, Kendari, Southeast Sulawesi, Indonesia

[3]Faculty of Health Sciences, Universitas Siliwangi, Jl. Siliwangi No. 24, Tasikmalaya, West Java, Indonesia

*wahabalwahab90@gmail.com

Abstract. This study examines the potential of biogas effluent as a source of electrical energy through microbial fuel cell (MFC) technology. The results showed that the one-room MFC produced a maximum voltage of 1302,2 mV, while the two-room MFC with $KMnO_4$ electrolyte reached 1084,3 mV with an electric current of 0,1 mA. Simultaneously, the two-chamber MFC with $K_2Cr_2O_4$ produced a maximum voltage of 1675,8 mV and an electric current of 0,99 mA. This potential, with electrical voltage values equivalent to commercial batteries, can improve the efficiency of electrical energy generation from organic waste and open up wider application opportunities in using renewable energy sources. Characterization of the MFC substrate showed a decrease in organic matter of biogas effluent with COD values of 500,96 mg/L and BOD of 300,23 mg/L. MFC bacteria from biogas waste were dominated by Gram-positive rod-shaped and Sarcina, except for one Gram-negative isolate of Spirillum.

Keywords: biogas, renewable energy, liquid waste, MFC

(Received 2023-12-25, Accepted 2024-01-03, Available Online by 2024-02-08)

1. Introduction

The high demand for electrical energy used in industry, transportation, and households in Indonesia is mainly from non-renewable fossil fuels [1]. By 2022, most of the electricity will come from coal (37,6%), natural gas (16,8%), fuel (33,4%) and renewable energy (NRE) (12,2%). This energy crisis

increases with society's and the economy's growth, making energy reserves increasingly depleted. Therefore, efforts are needed to develop renewable energy technologies, including from organic waste [2].

Biogas liquid waste, especially livestock manure sludge, is an organic waste that can be processed into an electrical energy source through microbial fuel cell (MFC) technology [3]. MFC is an electrochemical device that converts chemical energy from organic and inorganic materials into electrical energy through microbial metabolism, thus producing an environmentally friendly energy source [4]. This process generates electrons, which are captured and channelled through the electrodes, creating an electrical potential between the anode and cathode. An MFC system consists of two chambers, the anode and cathode chambers, where the substrate is oxidized and the oxidant (usually oxygen) is reduced [5].

In its development, various forms of organic matter can be used as MFC substrates, such as glucose [6] [4], starch [7] [8], fatty acids [9] [10], amino acids [11], proteins, leachate [9], and organic and inorganic sediments [1] [5]. MFCs have advantages in organic waste treatment and energy independence and are used for small-scale power generation. Despite the challenges of efficiency and durability of microorganisms, research continues for the optimization of this technology, presenting innovative solutions for sustainable energy needs [12].

Biogas effluent from cow dung has great potential as a substrate in MFC technology. The use of cow dung produces a maximum voltage of 0.723V at 50% concentration and is capable of powering a small LED lamp [13]. This research will focus on functionalizing cow dung biogas effluent into a renewable electrical energy source through MFC technology.

2. Methods

2.1. *Tools and Materials*

The tools used in this research are MFC reactor type two-compartment series, digital multimeter, N.Y.A ETERNA cable (1 x 2,5 mm), analytical balance, magnetic stirrer, beaker, measuring cup, erlenmeyer, measuring pipette, and pH meter. The materials used in the study were biogas liquid waste, distilled water, HCl 1 N, NaOH 1 N, standard solution of 5000 ppm glucose, $K_2Cr_2O_7$ 1 N, concentrated H_2SO_4, ferroin indicator, $[Fe(NH_4)_2(SO_4)_2]$ 0,2 N solution, NaCl solution, NH_4Cl, KH_2PO_4, $MgSO_4.5H_2O$, NaOH 40%, H_2SO_4 0,05 N, pure agar, KCl 1 M, KCl 0,01 M and buffer solution pH 7,0.

2.2. *Biogas Liquid Waste Sampling*

MFC substrate samples came from biogas liquid waste in Monapa Village, South Konawe. This effluent, the result of converting livestock manure into biogas, was stored in 5 L jerken and analyzed in the laboratory. This liquid waste becomes the main substrate in the MFC, with microorganisms playing a role in the oxidation of organic matter. Electrons generated during the process are captured by electrodes, generating electrical potential as an energy source to replace fuel oil.

2.3. *Characterization of Biogas Liquid Waste*

The characterization of biogas liquid waste refers to the research of Sudarman et al. [5] and Yao et al. [14], namely C-organic measurements, N-total measurements, COD (Chemical Oxygen Demand) measurements, BOD (Biological Oxygen Demand) measurements, pH measurements (H_2O and KCl) and electrical conductivity measurements.

2.4. *MFC Circuit Manufacturing*

The MFC electrodes used graphite rods from batteries. Previously, treatment was carried out by soaking the electrodes in 1 N HCl and 1 N NaOH, followed by rinsing with distilled water. The electrodes were then soaked again in distilled water before use. Each electrode was perforated and connected to the cable using epoxy and tested with a multimeter to ensure a successful connection. The MFC circuit followed the work of Zaeni et al [1]. Biogas effluent was put into a 500 mL beaker, with a

carbon electrode as the anode, connected to the cathode using a chemical electrolyte. The waste was allowed to stand for 20 hours, and then the voltage and current were measured using a multimeter on the MFC operated in dark conditions and at room temperature. Figure 1 shows the MFC circuit [5].

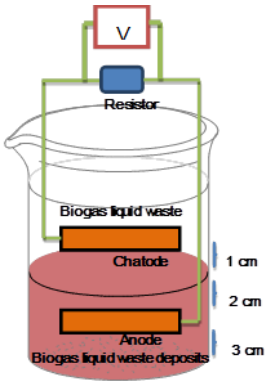


Figure 1. MFC circuit schematic

2.5. *Electrical Energy Measurement*

The electrical energy of the MFC system measured in this study is in the form of electric potential difference (voltage) and electric current. In the initial stage, measurements were taken for 20 hours and recorded every 2 hours. Measurements were made shortly after installation, which aimed to see the maximum open circuit voltage (open circuit voltage/Voc), which is a factor that affects an MFC in producing energy [1] [5].

2.6. *Isolation, Characterization, and Identification of MFC Bacteria*

The bacterial isolation stage consists of several steps: the preparation of liquid and solid media, bacterial inoculation, and bacterial isolation. The enrichment culture media used is modified APW (Alkaline Peptone Water) media [5]. Each litre of modified APW media contains 20 g NaCl, 0,77 g KCl, 0,25 g NH_4Cl, 0,1 g KH_2PO_4, 0,2 g $MgSO_4.7H_2O$ and 2,0 g $NaHCO_3$. The culture medium was then poured into rubber-stoppered tubes and autoclaved at 121°C. The solid media used for bacterial isolation was modified APW media added with pure agar (2%, wt/vol).

Bacterial inoculation in MFC involves taking 1 mL of biogas effluent and putting it into liquid media. The bacteria were incubated for two days at room temperature under dark conditions. After that, dilution and growth of bacteria on solid APW media were carried out. The growing bacterial colonies were isolated using the scratch cup method with an ose onto similar culture media. Characterization of bacterial isolates included observation of colony morphology, cell morphology (shape and Gram staining), and motility test. Bacterial identification was done manually with Bergey's Manual of Determinative Bacteriology.

3. Results and Discussion

3.1. *Characteristics of Biogas Liquid Waste*

After MFC, the biogas effluent changed from black to brown, indicating decreased organic matter levels [13]. Changes in organic matter levels before and after the use of MFC are recorded in Table 1.

Table 1. Characteristics of biogas waste liquid

Testing Parameters	Organic Matter Content	
	Before MFC	After MFC
C (%)	2,92	2,20
N (%)	2,01	1,85
C/N	1,45	1,19
COD (mg/L)	1001,92	500,96
BOD (mg/L)	805,54	300,23

DHL (dSm^{-1})	6,00	5,01
pH (H_2O)	7,24	7,67
pH (KCl)	6,87	7,54

Table 1 shows that the organic matter content of biogas liquid waste before MFC was 2,92%; after MFC, it decreased to 2,20%. Total nitrogen content before MFC was 2,01%, but after MFC, it decreased to 1,85%. Based on the carbon and nitrogen content analysis, the average C/N ratio of biogas liquid waste is 1,45. The value of the C/N ratio, which is smaller than 15, indicates the process of N mineralization in biogas liquid waste. N mineralization is the process of converting N-organic into N-organic by decomposing microbes. If the C/N ratio is greater than 30, N immobilization occurs; if it is between 15-30, mineralization is balanced with immobilization. The process of mineralization and immobilization of N in sewage is largely determined by the activity of sewage microorganisms, fungi, bacteria, and so on [1] [5].

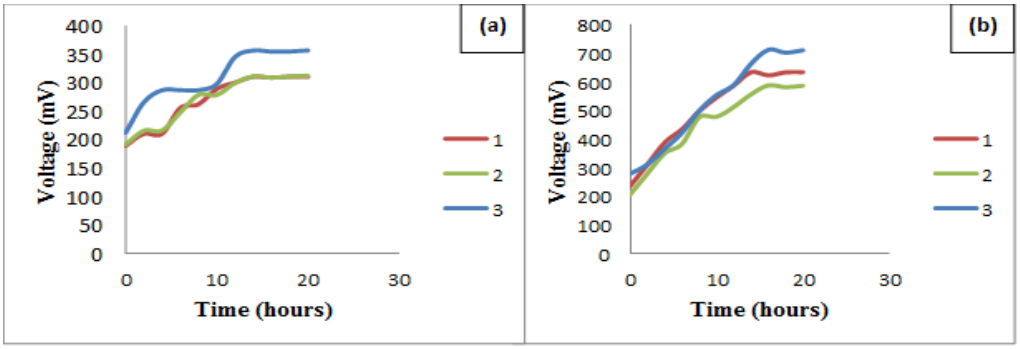
COD and BOD measurements support the production of electrical energy in MFCs. The higher the COD and BOD values, the greater the energy produced because microbes break down organic compounds, producing ions and electrons as a source of electrical energy. Table 1 shows a decrease in COD and BOD values after using biogas effluent as an MFC substrate, in line with the research of Nosek et al.[10]. Linh and Hai [11] also supported MFCs as BOD sensors, reducing the content of organic compounds in waste.

Table 1 shows the average pH of biogas effluent is around ±7. The pH range of 6,5-7,5 is optimal for nutrient availability, supporting the life of decomposing bacteria. At pH below 6,0, nutrient availability decreases, while pH above 8,0 increases the availability of certain elements. The electrical conductivity of biogas effluent is about six dSm^{-1}, providing good conductivity to MFC electrodes.

3.2. Open Circuit Voltage Measurement

Microbial Fuel Cell (MFC) is a device that converts chemical energy into electrical energy through the catalytic activity of microorganisms [15]. This bioelectrochemical system uses bacteria to convert organic matter into electrical energy [16]. MFCs can be made in various configurations, including one-chamber and two-chamber systems with $KMnO_4$ or $K_2Cr_2O_7$ electrolytes. One-room systems have the anode and cathode in the same space without a barrier, while two-room systems separate the anode and cathode with a salt bridge. The salt bridge maintains electrical charge neutrality, allowing the flow of electrons and ions, ensuring spontaneous redox reactions at the electrodes, resulting in a continuous electric current.

In this study, the open circuit voltage of the MFC system was measured for 20 hours to ensure the achievement of a steady state. Measurements were taken periodically every 2 hours to evaluate the changes in voltage over time. Figure 2(a) shows that the one-cell MFC system experienced an increase in voltage until it reached a steady state of about 310,6 mV. Meanwhile, the two-chamber MFC system with $KMnO_4$ and $K_2Cr_2O_7$ electrolytes reached a steady state at 312,3 mV and 357,1 mV, respectively. These results indicate that each MFC system can achieve stability in open circuit voltage after a certain time.



Description:
1. Single-chamber MFC system 2. Two-chamber MFC system using $KMnO_4$ electrolyte 3. Two-chamber MFC system using $K_2Cr_2O_7$ electrolyte

Figure 2. Comparison of time and voltage (open circuits voltage) generated by one-cell MFC system (a) and seven-cell and ten-cell MFC systems (b).

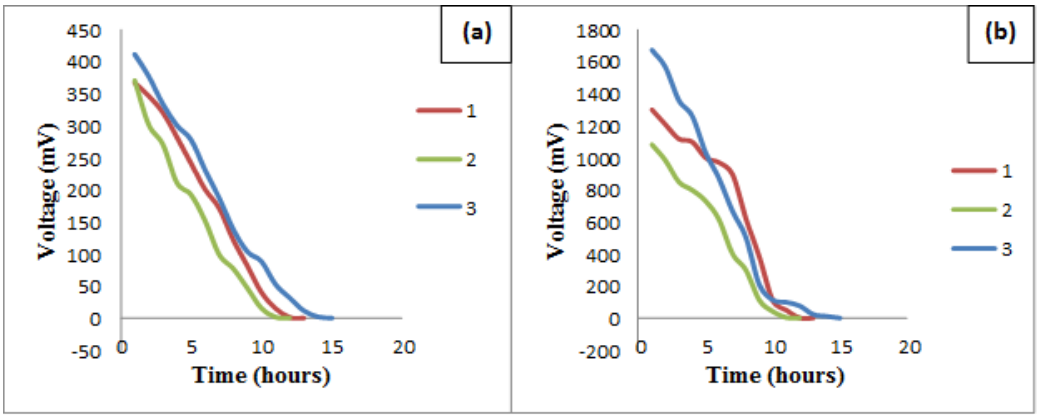
Figure 2 (b) shows that the open circuit voltage increases to 20 hours. In a one-room MFC system for ten cells, the open circuit voltage occurs in a steady state (fixed condition) at 634,1 mV. In a two-room MFC system using $KMnO_4$ electrolyte for seven cells, open circuit voltage occurs steady state (fixed condition) at 586,7 mV, while using $K_2Cr_2O_7$ electrolyte for ten cells open circuit voltage occurs steady state (fixed condition) at 711,4 mV In. The magnitude of the resulting voltage value is in line with the statement of Tiara et al. [17] that, in general, microbial activity has a peak point until the 18th hour. Microbial activity at the 18th hour is considered a log phase where microbes divide rapidly and constantly follow a logarithmic curve. The speed of growth in this phase is strongly influenced by the medium in which it grows, such as pH and nutrient content, as well as environmental conditions, including temperature and humidity. Microbes require more energy in this phase than in other phases [18].

3.3. MFC System Voltage and Current Measurements

The measurement of electrical energy in this study was carried out in parallel. A parallel circuit is an electrical circuit that can be installed between components in a row. According to Sudarman et al. [5], who took measurements with a parallel circuit, was able to increase the voltage value high enough but did not produce an electric current. The parallel circuit is expected to produce a large enough voltage and an electric current.

Based on the measurement results for the one-cell MFC system (Figure 3a.), the greatest voltage produced is from the dual chamber MFC system using $K_2Cr_2O_7$ electrolyte 412,1 mV. The measurement results of the voltage generated by one chamber MFC is 367,5 mV, and for the two-chamber MFC system using $KMnO_4$, it is 371,2 mV. The one-room MFC system assembled into ten cells (Figure 3b.) had a maximum voltage of 1302,2 mV, and no electric current was obtained. Measurement of two-chamber MFC using $KMnO_4$ electrolyte, which is assembled into seven cells, has a maximum voltage of 1084.3 mV and can produce an electric current of 0,01 mA, while measurement of MFC voltage using $K_2Cr_2O_7$ electrolyte produces a maximum voltage of 1675,8 mV and an electric current of 0,99 mA. The measurement results show that the highest voltage and current are produced in the two-chamber MFC system using the $K_2Cr_2O_7$ electrolyte.

In microbial metabolism, electrons (e^-) and protons (H^+) are generated, resulting in a potential difference. Electrons move from the anode to the cathode, and H^+ balances the difference through salt bridges. At the cathode, H^+ and e^- react with Cr(VI) metal ions to form Cr(III), which is called a reduction reaction. $K_2Cr_2O_7$, as a strong oxidizer, allows protons from the anode to be rapidly reduced to form H_2O, keeping the potential difference stable in the MFC system, according to the research of Xia et al. [19].



Description:
1. Single-chamber MFC system 2. Two-chamber MFC system using $KMnO_4$ electrolyte 3. Two-chamber MFC system using $K_2Cr_2O_7$ electrolyte

Figure 3. Comparison of time and voltage produced by one-cell MFC system (a) and seven-cell and ten-cell MFC systems (b).

The movement of the voltage measurement graph on the one-cell MFC system (Figure 3a.) and the ten-cell and seven-cell MFC systems (Figure 3b.) experienced a decrease in voltage from the first day of measurement to the last day of measurement. This is probably because the nutrients in the substrate have decreased, and there is no substrate addition, so this impacts microbial growth, which decreases. According to Goel [20], the amount of electrical energy produced by the MFC system can be influenced by microbial activity that utilizes the nutrients contained in the substrate. The more active the microbes are in breaking down organic matter in the substrate, the more free electrons will be produced in the anode chamber. The difference in the number of electrons produced between the anode and cathode chamber causes a potential difference reaction between the two, which a multimeter can detect.

In general, the results of this study have produced voltages that are relatively similar to commercial batteries but still have low amperage. Therefore, to be able to answer this challenge, a circuit is needed that can store energy from MFCs and run electronics at low potential (<500 mV) [21] [22]. This has been done by Lv et al. [23], who used a capacitor to store energy from microbial fuel cells and run sensors using the energy stored in the capacitor. Donovan et al. [21] have also tried to overcome these challenges by using a circuit design called Power Management System (PMS) to run a wireless sensor. The PMS design operates at low potential input regardless of the power generated by the SMFC. PMS starts running the wireless sensor when the SMFC potential reaches 320 mV and keeps running until the SMFC potential drops below 52 mV.

3.4. Characteristics of MFC Bacteria

Bacterial characterization research in biogas effluent using a numerical-phenetic approach. Phenotypic tests, gram painting, and motility tests were used. Identifying isolates aims to reveal specific traits of isolates, which is important for establishing taxonomic status. Four bacterial isolates with different characteristics were found, as noted in Table 2.

Biochemical characteristics of MFC bacterial motility testing showed that the bacterial isolates obtained were mostly nonmotile, classified as bacteria that cannot move or do not have flagella as a means of movement. Bacterial isolates that do not have a means of motion are isolates B_1, B_2, and B_4, while motile isolates with a means of motion are isolates B_3.

Table 2: Biochemical characteristics of MFC bacterial isolates

No.	Isolate	Salt Coloring	Motility Testing	Morphology Test		Description
				Cell Shape	Colony Shape	
1.	B_1	+	Nonmotil	Rod	Circular	Methanobacterium
2.	B_2	+	Nonmotil	Sarcina	Irregular	Methanosarcima
3.	B_3	-	Motil	Spirillum	Irregular	Methanospirillum
4.	B_4	+	Nonmotil	Rod	Circular	Methanobacterium

Gram staining is used to identify bacterial morphology and distinguish between Gram positive and Gram negative bacteria. Table 2 shows that isolates B_1, B_2, and B_4 from biogas effluent are Gram positive bacteria. Gram positive bacteria have cell walls that appear blue/purple under the microscope, due to lower lipid content and thicker peptidoglycan than Gram negative bacteria [24].

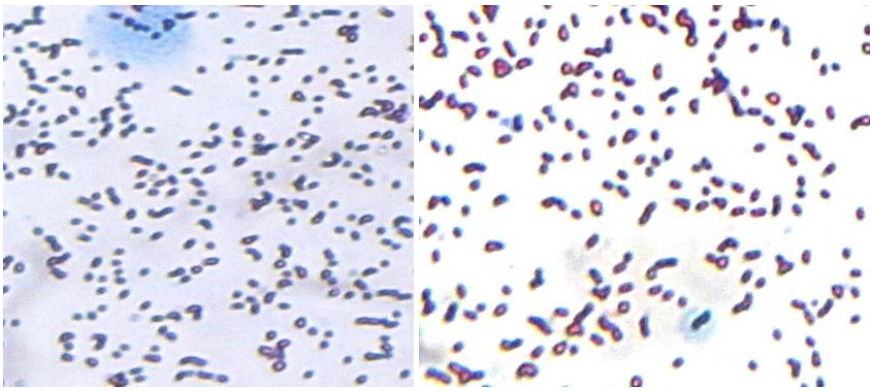


Figure 4. Appearance of bacteria under a microscope

From the specific properties presented in Table 2, it can be seen that the three isolates B_1 and B_4 are gram-positive bacteria with rod shape, B_2 is gram-positive with sarcina shape, and B_3 is gram-negative bacteria with spirillum shape. According to Bergey's Manual of Determinative Bacteriology, the characterisation test results in Table 2 can determine the genus of the bacterial isolates, namely isolates B_1 and B_4 are members of the genus Methanobacterium. Bacterial isolate B_2 is a member of the genus Methanosarcina, and B_3 belongs to the genus Methanospirillum.
Methanogen bacteria belong to one class of Archaebacteria in addition to halophilic and thermophilic by the group Archaebacteria are microorganisms that can survive in extreme areas such as waters with high salt content (halophil). Methanogenic bacteria are obligate anaerobes divided into three groups. Group I includes Methanobacterium and Methanobrevibacter, Group II includes Methanococcus, and Group III includes Methanospirillum and Methanosarcina genera. All are present in anaerobic freshwater environments such as sediments and in the digestive tract of animals such as cattle, sheep, and goats [25].

4. Conclusion

The seven-cell and ten-cell MFC systems, using $K_2Cr_2O_7$ solution, produced a maximum electrical energy of 1675,8 mV and an electrical current of 0,99 mA. After 20 days, the biogas effluent underwent changes with a decrease in organic matter and nitrogen content, resulting in a C/N ratio of 1,19. The effluent pH increased on average to 7,67 (pH H_2O) and 7,54 (pH KCl), with an electrical conductivity of 5,01 dSm^{-1}. The research demonstrates the potential of MFCs as a renewable energy source, a contribution to energy diversification, and the benefits of efficient waste management.

References

[1] A. Zaeni, P. E. Susilowati, Alwahab, and L. O. Ahmad, "Renewable energy from sediment microbial fuel cell technology from Kendari Bay swamp sediments," *AIP Conf. Proc.*, vol. 2237, no. June, 2020, doi: 10.1063/5.0011271.

[2] IESR, "Indonesia Energy Transition Outlook 2023: Tracking Progress of Energy Transition in Indonesia: Pursuing Energy Security in the Time of Transition," p. Please cite this report as: IESR (2022). Indonesia, 2022, [Online]. Available: www.irena.org

[3] B. Hülsemann, L. Zhou, W. Merkle, J. Hassa, J. Müller, and H. Oechsner, "Biomethane potential test: Influence of inoculum and the digestion system," *Appl. Sci.*, vol. 10, no. 7, 2020, doi: 10.3390/app10072589.

[4] M. Christwardana, D. Frattini, G. Accardo, S. P. Yoon, and Y. Kwon, "Optimization of glucose concentration and glucose/yeast ratio in yeast microbial fuel cell using response surface methodology approach," *J. Power Sources*, vol. 402, no. June, pp. 402–412, 2018, doi: 10.1016/j.jpowsour.2018.09.068.

[5] R. Sudarman, A. Zaeni, I. Usman, Alwahab, and I. Hidayatulloh, "Optimization of electricity generation from marine sediment of Kendari Bay using stacked sediment microbial fuel cell," *J. Phys. Conf. Ser.*, vol. 1450, no. 1, 2020, doi: 10.1088/1742-6596/1450/1/012111.

[6] C. Montoya-vallejo, J. Omar, G. Posada, and J. Carlos, "Effect of Glucose and Methylene Blue in Microbial Fuel Cells Using E. coli," *Energies*, vol. 16, no. 7901, 2023, doi: https://doi.org/10.3390/en16237901.

[7] C. Chen, L. Zong, J. Wang, and J. Xie, "Microfibrillated cellulose reinforced starch/polyvinyl alcohol antimicrobial active films with controlled release behavior of cinnamaldehyde," *Carbohydr. Polym.*, vol. 272, no. April, p. 118448, 2021, doi: 10.1016/j.carbpol.2021.118448.

[8] N. V. Patil and A. N. Netravali, "Microfibrillated cellulose-reinforced nonedible starch-based thermoset biocomposites," *J. Appl. Polym. Sci.*, vol. 133, no. 45, pp. 2–11, 2016, doi: 10.1002/app.43803.

[9] D. Nosek, O. Samsel, T. Pokój, and A. Cydzik-Kwiatkowska, "Waste volatile fatty acids as a good electron donor in microbial fuel cell with the iron-modified anode," *Int. J. Environ. Sci. Technol.*, vol. 20, no. 12, pp. 13021–13032, 2023, doi: 10.1007/s13762-023-04850-8.

[10] D. Nosek and A. Cydzik-Kwiatkowska, “Microbial structure and energy generation in microbial fuel cells powered with waste anaerobic digestate,” *Energies*, vol. 13, no. 18, 2020, doi: 10.3390/en13184712.
[11] N. Thi My Linh and P. The Hai, “An in-depth study of substrate effect on the performance of BOD sensor type microbial fuel cell,” *Vietnam J. Biotechnol.*, vol. 20, no. 1, pp. 173–185, 2022, doi: 10.15625/1811-4989/16056.
[12] J. Wang, K. Ren, Y. Zhu, J. Huang, and S. Liu, “A Review of Recent Advances in Microbial Fuel Cells: Preparation, Operation, and Application,” *BioTech*, vol. 11, no. 4, 2022, doi: 10.3390/biotech11040044.
[13] P. Anand, A. Shaheen, A. Masroor, A. S. Suhail, and K. Aisha, “Design and fabrication of microbial fuel cell using cow manure for power generation. Sci.Int.(Lahore),27(5),4235-4238,” vol. 27, no. 5, pp. 4235–4238, 2006.
[14] H. Yao, J. Xiao, and X. Tang, “Microbial Fuel Cell-Based Organic Matter Sensors: Principles, Structures and Applications,” *Bioengineering*, vol. 10, no. 8, 2023, doi: 10.3390/bioengineering10080886.
[15] A. S. Yaakop, A. Ahmad, F. Hussain, S. E. Oh, M. B. Alshammari, and R. Chauhan, “Domestic Organic Waste: A Potential Source to Produce the Energy via a Single-Chamber Microbial Fuel Cell,” *Int. J. Chem. Eng.*, vol. 2023, 2023, doi: 10.1155/2023/2425735.
[16] G. Kaur *et al.*, “Management of Cattle Dung and Novel Bioelectricity Generation Using Microbial Fuel Cells: An Ingenious Experimental Approach,” *Int. J. Chem. Eng.*, vol. 2021, 2021, doi: 10.1155/2021/5536221.
[17] A. Info, “Jurnal Bahan Alam Terbarukan Development of Microbial Fuel Cell in Tofu Liquid Waste in Producing Clean,” vol. 12, no. 200, pp. 97–104, 2023.
[18] P. Dessì, E. Porca, J. Haavisto, A. M. Lakaniemi, G. Collins, and P. N. L. Lens, “Composition and role of the attached and planktonic microbial communities in mesophilic and thermophilic xylose-fed microbial fuel cells,” *RSC Adv.*, vol. 8, no. 6, pp. 3069–3080, 2018, doi: 10.1039/c7ra12316g.
[19] C. Xia, D. Zhang, W. Pedrycz, Y. Zhu, and Y. Guo, “Models for Microbial Fuel Cells: A critical review,” *J. Power Sources*, vol. 373, no. July 2017, pp. 119–131, 2018, doi: 10.1016/j.jpowsour.2017.11.001.
[20] S. Goel, “From waste to watts in micro-devices: Review on development of Membraned and Membraneless Microfluidic Microbial Fuel Cell,” *Appl. Mater. Today*, vol. 11, pp. 270–279, 2018, doi: 10.1016/j.apmt.2018.03.005.
[21] C. Donovan, A. Dewan, D. Heo, and H. Beyenal, “Batteryless, wireless sensor powered by a sediment microbial fuel cell,” *Environ. Sci. Technol.*, vol. 42, no. 22, pp. 8591–8596, 2008, doi: 10.1021/es801763g.
[22] S. Gupta *et al.*, “The race between classical microbial fuel cells, sediment-microbial fuel cells, plant-microbial fuel cells, and constructed wetlands-microbial fuel cells: Applications and technology readiness level,” *Sci. Total Environ.*, vol. 879, no. January, 2023, doi: 10.1016/j.scitotenv.2023.162757.
[23] C. Lv, J. Liu, Y. Zhang, W. Lei, R. Cao, and G. Lv, “Reliability Modeling for Metallized Film Capacitors Based on Time-Varying Stress Mission Profile and Aging of ESR,” *IEEE J. Emerg. Sel. Top. Power Electron.*, vol. 9, no. 4, pp. 4311–4319, 2021, doi: 10.1109/JESTPE.2020.3001604.
[24] N. Nano, “Science and Technology Information 科技信息 Science and Technology Information,” vol. 3, no. 1, pp. 28–31, 2016.
[25] M. R. Swain, R. C. Ray, and C. S. Nautiyal, “Biocontrol efficacy of Bacillus subtilis strains isolated from cow dung against postharvest yam (Dioscorea rotundata L.) pathogens,” *Curr. Microbiol.*, vol. 57, no. 5, pp. 407–411, 2008, doi: 10.1007/s00284-008-9213-x.

Advance Sustainable Science, Engineering and Technology (ASSET)
Vol. 6, No.2, April 2024, pp. 0240203-01 ~ 0240203-08
ISSN: 2715-4211 DOI: https://doi.org/10.26877/asset.v6i2.18228

Comparing Optimizer Strategies For Enhancing Emotion Classification In IndoBERT Models

Julius Immanuel Theo Krisna[1*], Ardytha Luthfiarta[1], Leno Dwi Cahya[1], Sri Winarno[1], Adhitya Nugraha[1]

[1]Faculty of Computer Science, Universitas Dian Nuswantoro Semarang, Jl. Imam Bonjol No. 207 Semarang, Central Java, Indonesia

*juliusimmanueltheo@gmail.com

Abstract. Emotions are one of the reactions of human when they receive physical or verbal action. Every human action is based on emotion. Every opinion expressed in the comments column also contains the author's emotions. This research aims to classify five emotions, Marah, Takut, Senang, Cinta, and Sedih and evaluate the performance of three commonly used optimizer, Adam, RMSProp, and Nadam. The processed data used IndoBERT model for Indonesian text classification. The research purpose to search the best optimizer for text classification. The result shows classification used Adam Optimizer 90,21%, RMSProp Optimizer 82.11, and Nadam Optimizer 88.61%. The Adam optimizer applied to the IndoBERT model yielded the best results. This shows a significant improvement from previous studies, which had emotion classification.

Keywords: NLP, IndoBERT, Optimizer, Adam, RMSProp, Nadam

(Received 2024-01-20, Accepted 2024-01-31, Available Online by 2024-02-08)

1. Introduction

Emotions are one of the human reactions when receiving physical or verbal actions. Emotions can be shown through facial expressions or actions.[1] Emotions are experiences that humans experience consciously to characterize a psychological state that is important as part of human nature, such as joy, anger, love, sadness, fear, and so on.[2] Every human action is based on emotions. Every opinion expressed in the comments column also contains the author's emotions. Dissemination of information can influence public opinion. The emotional content in this opinion is interesting to analyze in research, especially the emotional patterns contained in this opinion. Text classification aims to analyze, process, and extract information contained in text. In emotion classification, a comment will be taken to obtain the information contained in it so that we can find out the emotion in the comment. [3]

Many companies currently use comments from various social media in the form of Instagram, Twitter, and Facebook.[4] Through this, they can understand customers' feelings about the company. This opinion processing usually uses sentiment analysis which has developed in various fields of business, government, and organizations.[5] Natural Language Processing (NLP) plays an important role in the field of Artificial Intelligence (AI).

Natural Language Processing enables machines to understand, interpret, and respond to human language, thereby facilitating humancomputer interactions. NLP is tightly integrated with machine learning (ML) and AI, enabling the development of voicecontrolled systems, language translation, sentiment analysis, and a variety of other applications. The combination of NLP, AI, and ML drives the automation of data analysis processes and changes the way we interact with technology, making advanced tools more accessible and user-friendly. [6] The development of artificial intelligence technology has been enormous in recent years due to the use of the Internet, big data, the Internet of Things, and the use of very massive processing power.[7], [8]

Nowadays pre-trained language models have helped increase the sophistication in many areas of NLP (Natural Language Processing). Currently, many machine learning models are being developed for sentiment analysis. This includes conventional models such as SVM (Support Vector Machine), NBC (Naïve Bayes Classifier), and currently the use of deep learning models based on Neural Networks. One of the developments in the deep learning model is Transformer, including Bidirectional Encoder Representations from Transformers (BERT). [9], [10], [11]

In recent years Bidirectional Encoder Representations from Transformers (BERT) combines the representation of words and sentences in a Transformer on a large scale. Determining the right model with a dataset is a way to overcome large and heavy computations.[12] In BERT there are various types of models such as IndoBERT, ALBERT, RoBERTa, and many more. Indonesian is the 10th most widely used language. Because of its wide use, NLP practitioners use the IndoBERT model to organize existing language resources.[1] IndoBERT is a transformer model by adapts BERT itself but is drilled purely as a masked language model drilled using huggingface by following the BERT base configuration.[13] BERT models are a development of Transformer by adapting the dataset used by the user. The use of models that match the language in the dataset also reduces the computational complexity.[4] Optimization is one aspect of improving the performance of the model used.[14]

This research aims to implement the Indonesian Transformer model, namely the IndoBERT model, to compare the three optimization methods Adam (Adaptive Moment Estimation), Nadam (Nesterov-accelerated Adaptive Moment Estimation), and RMSProp (Root Mean Square Propagation). Optimization plays an important role in managing the accuracy by adjusting the learning rate of the model during the training process to get optimal prediction results.

Table 1. Literature Study

Author	Year	Method	Result or Finding
Bagus [1]	2022	Comparison BERT Uncased and IndoBERT with Adam Optimizer	Classifying emotions using BERT and Adam, resulting in an accuracy value of 90% in BERT Uncased and producing result 81% accuracy in using the IndoBERT.
Hulliyah, Rayya, Bakar [15]	2022	IndoBERT with Adam Optimizer for develop Chatbot	The use of optimizers in Natural Language Processing using IndoBERT with the Adam optimizer to develop chatbots for emotion classification obtained accuracy, F1 score, recall, and precision values of 89%, 89%, 89%, and 90% on train data. To validate the data, the accuracy, F1 score, recall, and precision values were 70%, 71%, 70%, and 72%.
Wijaya [16]	2021	IndoBERT with Adam Optimizer	In implementing IndoBERT with another Adam optimizer Disabled News classification, getting an accuracy value of 0.853, recall 0.853, precision 0.872, and F1 0.853.
Cahya [12]	2024	Handling Imbalaced Dataset using IndoBERT	Experiment with AdamW optimizer the result is Augmentation technique performance by up to 20%, with accuracy 78%, precision 85%, recall 82%, and an

		Model	F1-score of 83%. With SMOTE technique, the evaluation results achieve accuracy to a high 82% with precision at 87%, recall at 85%, and an F1-score of 86%.
Nugroho, Bachtiar [17]	2021	Text-Based Emotion Recognition in Indonesian Tweet using BERT	Experiment use Adam optimizer for BERT-Base got 0,6478 accuracy and IndoBERT-Base got 0,7673 accuracy.

Based on the provided literature study, it is possible to utilize IndoBERT-Base for classifying emotion from Dataset Twitter. The usage of IndoBERT for emotion classification is still relatively uncommon. However, in some studies, the effectiveness of IndoBERT is reported to be suboptimal compared to BERT-Uncased. [1] Several research works have explored the potential of optimizing IndoBERT's performance, with findings suggesting that the use of the Adam Optimizer can yield improved results. [1], [15], [16], [17] IndoBERT-Base remains relevant for classification purposes to date.

This research endeavors to surpass previous studies by employing advanced methods and algorithms. The investigation employs IndoBERT-Base with a focus on emotion classification, utilizing a dataset with five labels yaitu Marah (Angry), Takut (Fear), Senang (Joy), Cinta (Love), and Sedih (Sad) from 9,480 Tweets. The dataset is larger than that used in [17]. Additionally, three different optimizers are employed to identify the most effective one for achieving optimal classification results. The primary goal of this study is to specifically determine the best optimizer for emotion classification using pre-trained IndoBERT-Base.

2. Methods

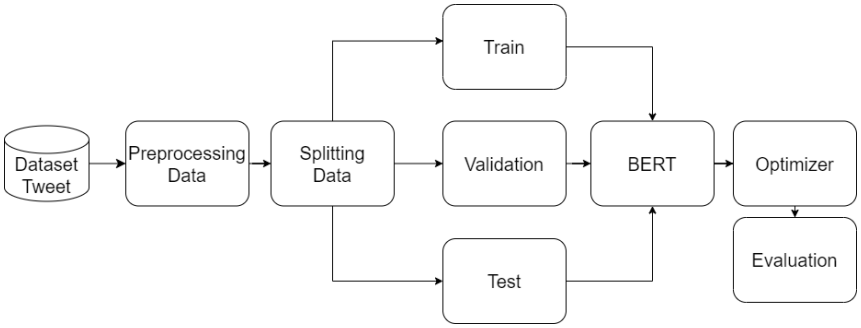


Figure 1. Research Methods

The research begins by carrying out a dataset search stage which can be illustrated in Figure 1. The dataset has been adjusted through the Preprocessing Stage for model training. Next, the dataset will be divided into 3 parts: Train Data, Validation Data, and Test Data. In the training process, the model is processed in Data Train, and its performance is assessed using various evaluation metrics, including graphical accuracy, graphical loss, confusion matrix, and classification results. Evaluation of this Matrix can provide new insights into the effectiveness of the model and its ability to classify and predict results correctly.

2.1. Dataset

In this research, we use the Emotion Dataset from public opinion from Twitter. This research uses a multi-label emotion dataset in Indonesian. This dataset has 5 labels, namely Marah (Angry), Takut (Fear), Senang (Joy), Cinta (Love), and Sedih (Sad). This dataset has a total of 9,480 tweets along with labels.

Table 2. Dataset Emotion

Tweet	**Label**
bikin gue emosi aja	Marah
takut bgt jujur kalo yangti udh turun tangan	Takut
waa enak tuu, semoga libur mu bahagia yaa!~	Senang

suka nih netijen yang kayak gini :) otak" detektif gw like banget	Cinta
mandi sore, soalnya mo les. sedih malem minggu malah les	Sedih

2.2. Text Preprocessing

This research uses several preprocessing, including: case folding such as lower case to change capital text to lower case and delimiters to remove parts that can influence results such as tags, urls, emoji, punctuation and whitespace. Next, the data is given a tokenizer provided in the BERT model. In addition, data that null will be removed.

2.3. Modelling IndoBERT

In this research, classification is carried out using the IndoBERT model, which is a development of BERT (BiDirectional Encoder Representations Transformers). This difference is because the architecture of this model focuses on processing and training data in Indonesian based on the Transformer architecture. In the Transformer architecture, there are two parts, namely the decoder and the encoder. Specifically for BERT, it only uses the Encoder architecture.[18]

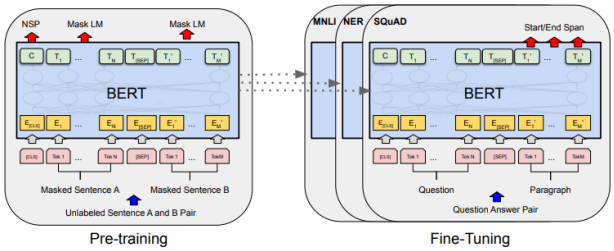


Figure 2. BERT architecture [18]

The BERT model architecture is a multi-layer bidirectional Transformer. There are two stages carried out in BERT. There are two stages in BERT, namely the pre-training stage and the fine-tuning stage. In the pre-training stage, the BERT model architecture is drilled using unlabeled data and trained with different tasks, while for the fine-tuning stage, the BERT model is first initialized with pre-drilled parameters, and all parameters are tuned using labeled data.[13]

In the pretrain stage, the model is drilled with a labeled dataset. For the IndoBERT model architecture, a large dataset was collected, consisting of about four billion words and about 250 million sentences from the Indonesian text collection. This dataset includes news texts from various sources such as local online, social media, Wikipedia, online articles, subtitle texts from video recordings, and a parallel data set known as Indo4B. Indo4B includes formal word data, informal word data and casual words in Indonesian.[12]

2.4. Optimizer

2.4.1. Adam (Adaptive Moment Estimation)

Adam is one of the optimizations that is widely used in training models today. This optimization can be used in machine learning or deep learning. Adam is a gradual optimization algorithm that uses an adaptive learning rate. The Adam optimizer combines the first and second moments of the gradient to update the parameters.

$$m_t = \beta_1 . m_{t-1} + (1 - \beta_1) . g_t \quad (1)$$

$$v_t = \beta_2 . v_{t-1} + (1 - \beta_1) . g_t^2 \quad (2)$$

$$\widehat{m_1} = \frac{m_1}{1-\beta_1^t} \quad (3)$$

$$\widehat{m_2} = \frac{m_2}{1-\beta_2^t} \quad (4)$$

$$p = \frac{p - lr \times \widehat{m_1}}{\sqrt{\widehat{m_2}} + e} \quad (5)$$

In this formula, m is the momentum, β is the repetition parameter, g is the gradient of the parameter loss function, p is the parameter, t is the iteration, lr is the learning rate, and epsilon is a small number to avoid divide it by zero.

2.4.2. RMSProp (Root Mean Square Propagation)

RMSPop is the default optimization algorithm used in some neural network models. These algorithms have adjustments for the problem of decreasing the learning rate too quickly.

$$\theta_{t+1} = \theta_t - \frac{\eta}{\sqrt{E[g^2]_t + \epsilon}} \cdot g_t \quad (6)$$

In this formula $\theta_t$is the parameter at iteration time (t), $g_t$is the gradient parameter at iteration time (t), $E[g^2]_t$is the storage of the average square of the gradient at iteration time (t), η is the learning rate, and ϵ is the epsilon is a small number to avoid divide it by zero.

2.4.3. Nadam (Nesterov-accelerated Adaptive Moment Estimation)

Nadam is a development of Adam. Nadam is a combination of the Nesterov Accelerated Gradient (NAG) optimizer with Adam (Adaptive Moment Estimation).

$$\theta_{t+1} = \theta_t - \alpha \left(\beta_1 m_t \frac{(1-\beta_1)\nabla f(\theta_t)}{1-\beta_1^t}\right) \cdot \left(\frac{\beta_2 \sqrt{v_t}}{\sqrt{1-\beta_2^t}} + \epsilon\right)^{-1} \quad (7)$$

In this formula $\theta_t$is the parameter at iteration time (t), α is the learning rate, β_1 and $\beta_2$are the decay factors for gradient and gradient square respectively. m_t is the average change in gradient, v_t is the average change in the squared gradient, and ϵ is a small value that prevents division by zero.

2.5. Evaluation

In this research, evaluation was carried out with three models, namely Adam, RMSProp, and Nadam. Evaluation has been carried out by comparing several optimizers with a confusion matrix. The results of the confusion matrix will be compared and explained to determine the performance of each optimizer. The performance will be compared in terms of precision, recall, f1-score, and accuracy.

3. Results and Discussion

3.1. Dataset

In this research, we use the Emotion Dataset from public opinion from Twitter. This dataset contains two columns containing tweets and labels. This research uses a multi-label emotion dataset in Indonesian. This dataset has five labels, namely Marah (Angry), Takut (Fear), Senang (Joy), Cinta (Love), and Sedih (Sad). This dataset has a total of 9,480 tweets along with labels.

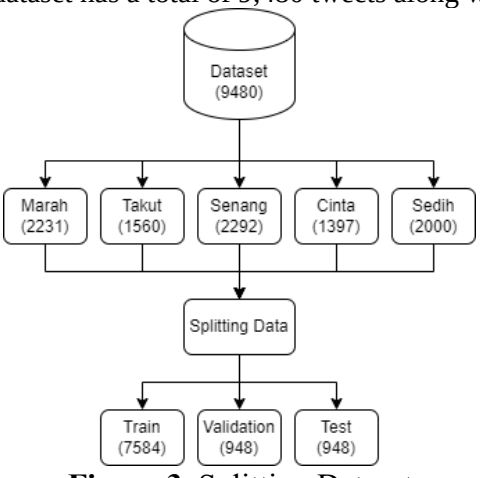


Figure 3. Splitting Dataset

This research uses a dataset of 9,480 text data from the Emotion Dataset in Indonesian which is divided into 5 labels love 1397, anger 2231, sadness 2000, happy 2292, fear 1560. After being combined the dataset is divid-ed into three parts for train, validation, and test of 8:1:1. The amount of data divided into each is 7580 for train data, 948 for validation data, and 948 for test data.

3.2. Percetage Training Result

Emotion classification experiments were carried out using the IndoBERT model. The dataset comes from Tweets that have been labeled. The existing dataset is classified into five labels. The labels consist

of 'Love', 'Angry', 'Sad', 'Happy', and 'Fear'. The control variables in training used in this model consist of thje dataset, IndoBERT model, learning rate 1e-6, and batch size 8. In the experimental process, 25 epochs were used by applying an early stop and patient of 5 epochs.

Table 3. Comparison Accuracy and Confusion Matrix

Optimizer	Accuracy		Average Confusion Matrix		
	Train	Test	Precision	Recall	F1-Score
Adam	90,21%	80.59%	81%	81%	81%
RMSProp	82.11%	78.37%	79%	79%	79%
Nadam	88.61%	80.90%	80%	80%	80%

In this experiment, the highest accuracy results were obtained on the Adam optimizer. On the other hand, the results of the RMSProp optimizer are more stable compared to other optimizers. In the Average Confusion Matrix results, each optimizer does not have a striking difference because the distance between each optimizer is not far.

3.3. Graph Training Result

In the research, training was carried out using the IndoBERT model. The existing dataset is classified into five labels. The control variables in the training used in this model consist of the dataset, IndoBERT model, learning rate 1e-6, and batch size 8. In the process used, 25 epoch training is applied by applying an early stop and patient of 5 epochs.

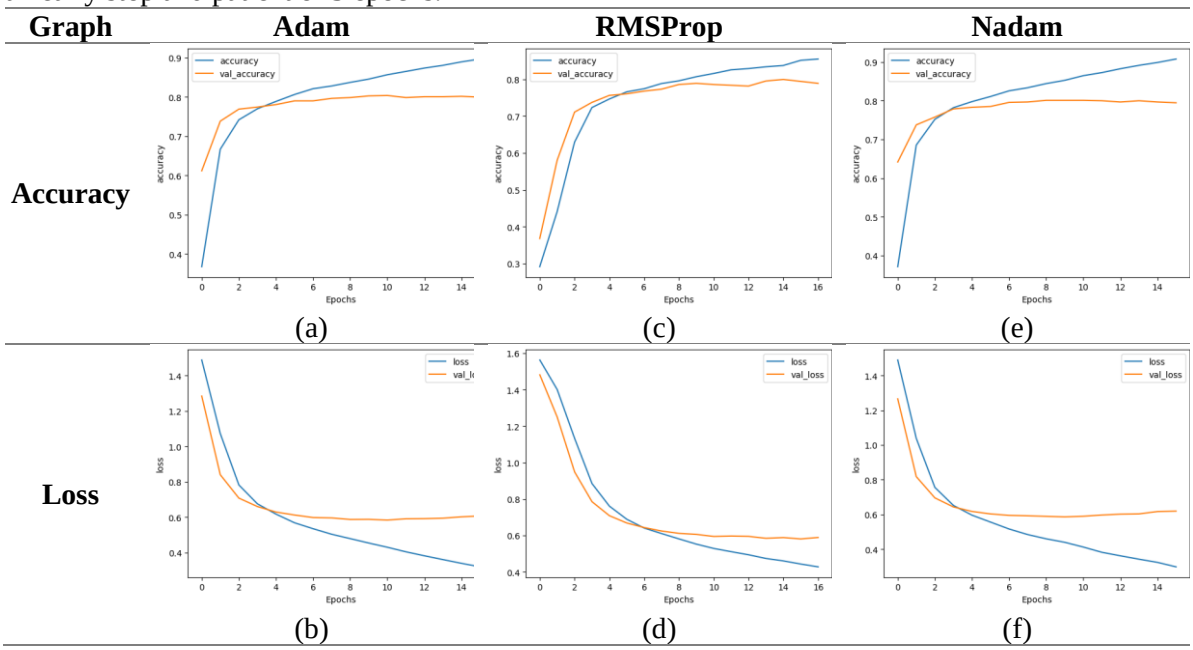


Figure 4. Graph Result (a) accuracy (b) loss

In Figure 4 (a) and (b), the graph of train and validation data using the Adam optimizer (a) and (b) shows signs of over-fitting. The training results show that with the IndoBERT optimizer model Adam obtained quite high accuracy results compared to the others, but the results obtained showed quite a large difference between train and validation accuracy. In this research, the Adam optimizer can show the potential for overfitting, where the model learns too many details from the training data that may not be common and cannot be applied to new data.

Figure 4 (c) and (d), shows that in the graph using the RMSProp optimizer, the results obtained are quite good because the test results obtained can look stable. Although the obtained RMSProp Optimizer training accuracy is lower than Adam's, the difference between training and validation accuracy is smaller. This graph shows the model tends to be more general and perhaps better at handling new data.

This optimizer may be more resistant to overfitting. The RMSProp optimizer as the default optimizer in the IndoBERT model has a striking difference between train and validation accuracy which is smaller. This shows that the RMSProp optimizer tends to be more stable in processing train, validation, and test data.

Figure 4 (e) and (f), shows that the graph using the Nadam Optimizer in this model has the highest accuracy on train, validation, and test data, and the difference between training and validation accuracy is very large. In this re-search, the Adam optimizer can show the potential for overfitting, where the model is too large. learning de-tails from training data that may be unfamiliar and not applicable to new data.

Data overfitting arises because the distance between the training and validation graphs is quite large. As in the Tabel 3, Adam and Nadam Optimizer graph the train graph is growing but the validation graph is stagnant. The results from RMSProp get a fairly close distance difference between the training and validation graphs. If a model is trained too specifically on the training data, it will likely perform poorly on new data because the model has "memorized" patterns that are not generally patterns that can be applied to new data.

3.4. Anomaly Result

In this experiment, several reasons resulted in the data being difficult for the model to classify. Data in the form of sentences can produce different results when processed. Sentences that contain two or more emotions in them can produce different emotions in each process. On the other hand, in the preprocessing process of removing punctuation marks, several sentences have different outputs because semantically the sentences have different meanings.

4. Conclusion

Based on the results obtained, determining the right optimizer is done by conducting experiments according to the desired aspects. This research aims to classify five emotions, Marah, Takut, Senang, Cinta, and Sedih and evaluate the performance of three commonly used optimizer, Adam, RMSProp, and Nadam. The processed data used IndoBERT model for Indonesian text classification. The research purpose to search the best optimizer for text classification. This research uses control variables in the training used in this model consisting of dataset, IndoBERT model, learning rate 1e-6, and batch size 8. In the training process, 25 epochs are used by applying an early stop and patient of 5 epochs.

The result shows classification used Adam Optimizer 90,21%, RMSProp Optimizer 82.11, and Nadam Optimizer 88.61%. The Adam optimizer applied to the IndoBERT model yielded the best results. This shows a significant improvement from previous studies, which had emotion classification. There are still many ways to improve the performance of Indonesian Text Processing but Indonesian text data processing resources are still lacking and still need to be developed further. Emotion classification can be further developed with multilabels by processing each emotion in one data.

References

[1] A. Tusa Bagus, “KLASIFIKASI EMOSI PADA TEKS MENGGUNAKAN METODE DEEP LEARNING,” 2022.

[2] R. D. Handayani, K. Kusrini, and H. Al Fatta, “Perbandingan Fitur Ekstraksi Untuk Klasifikasi Emosi Pada Sosial Media,” *Jurnal Ilmiah SINUS*, vol. 18, no. 2, p. 21, Jul. 2020, doi: 10.30646/sinus.v18i2.457.

[3] A. S. Aribowo and S. Khomsah, “Implementation Of Text Mining For Emotion Detection Using The Lexicon Method (Case Study: Tweets About Covid-19) Implementasi Text Mining Untuk Deteksi Emosi Menggunakan Metode Leksikon (Studi Kasus: Twit Tentang Covid-19),” *Jurnal Informatika dan Teknologi Informasi*, vol. 18, no. 1, pp. 49–60, 2021, doi: 10.31515/telematika.v18i1.4341.

[4] T. Shaik, X. Tao, C. Dann, H. Xie, Y. Li, and L. Galligan, “Sentiment analysis and opinion mining on educational data: A survey,” *Natural Language Processing Journal*, vol. 2, p. 100003, Mar. 2023, doi: 10.1016/j.nlp.2022.100003.

[5] W. Zhang, X. Li, Y. Deng, L. Bing, and W. Lam, "A Survey on Aspect-Based Sentiment Analysis: Tasks, Methods, and Challenges," *IEEE Trans Knowl Data Eng*, vol. 35, no. 11, pp. 11019–11038, Nov. 2023, doi: 10.1109/TKDE.2022.3230975.

[6] M. N. Hoda, Bharati Vidyapeeth's Institute of Computers Applications and Management Delhi, and Institute of Electrical and Electronics Engineers Delhi Section, *NLP & AI Speech Recognition: An Analytical Review*. 2023.

[7] Z. Li, Y. Wang, Y. Ji, and W. Yang, "A Survey of the Development of Artificial Intelligence Technology," in *2020 3rd International Conference on Unmanned Systems (ICUS)*, IEEE, Nov. 2020, pp. 1126–1129. doi: 10.1109/ICUS50048.2020.9274952.

[8] W. Hongyuan and D. Mingxing, "OVERVIEW OF THE DEVELOPMENT OF ARTIFICIAL INTELLIGENCE TECHNOLOGY," *Int J Res Eng Technol*, vol. 07, no. 08, pp. 92–95, Aug. 2018, doi: 10.15623/ijret.2018.0708011.

[9] M. Wankhade, A. C. S. Rao, and C. Kulkarni, "A survey on sentiment analysis methods, applications, and challenges," *Artif Intell Rev*, vol. 55, no. 7, pp. 5731–5780, Oct. 2022, doi: 10.1007/s10462-022-10144-1.

[10] Y. Yanfi, Y. Heryadi, L. Lukas, W. Suparta, and Y. Arifin, "Sentiment Analysis of User Review on Indonesian Food and Beverage Group using Machine Learning Techniques," in *2022 IEEE Creative Communication and Innovative Technology (ICCIT)*, IEEE, Nov. 2022, pp. 1–5. doi: 10.1109/ICCIT55355.2022.10118707.

[11] S. Saadah, Kaenova Mahendra Auditama, Ananda Affan Fattahila, Fendi Irfan Amorokhman, Annisa Aditsania, and Aniq Atiqi Rohmawati, "Implementation of BERT, IndoBERT, and CNN-LSTM in Classifying Public Opinion about COVID-19 Vaccine in Indonesia," *Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi)*, vol. 6, no. 4, pp. 648–655, Aug. 2022, doi: 10.29207/resti.v6i4.4215.

[12] L. D. Cahya, A. Luthfiarta, J. I. T. Krisna, S. Winarno, and A. Nugraha, "Improving Multi-label Classification Performance on Imbalanced Datasets Through SMOTE Technique and Data Augmentation Using IndoBERT Model," *Jurnal Nasional Teknologi dan Sistem Informasi*, vol. 9, no. 3, pp. 290–298, Jan. 2024, doi: 10.25077/TEKNOSI.v9i3.2023.290-298.

[13] F. Koto, A. Rahimi, J. H. Lau, and T. Baldwin, "IndoLEM and IndoBERT: A Benchmark Dataset and Pre-trained Language Model for Indonesian NLP," Nov. 2020, [Online]. Available: http://arxiv.org/abs/2011.00677

[14] P. Langgeng, W. E. Putra, M. Naufal, and E. Y. Hidayat, "A Comparative Study of MobileNet Architecture Optimizer for Crowd Prediction," *Semarang 123 Jl. Imam Bonjol No*, vol. 8, no. 3, p. 50131, 2023.

[15] K. Hulliyah, F. Rayyan, and N. S. A. A. Bakar, "Development Of A Chatbot For The Online Application Telegram Chat With An Approach To The Emotion Classification Text Using The Indobert-Lite Method," in *2022 4th International Conference on Cybernetics and Intelligent System (ICORIS)*, IEEE, Oct. 2022, pp. 1–4. doi: 10.1109/ICORIS56080.2022.10031483.

[16] A. Wijaya, "Implementasi Model Neural Network IndoBERT untuk Klasifikasi Berita Ddifabel," Universitas Multimedia Nusantara, Jakarta, 2021.

[17] K. S. Nugroho and F. A. Bachtiar, "Text-Based Emotion Recognition in Indonesian Tweet using BERT," in *2021 4th International Seminar on Research of Information Technology and Intelligent Systems (ISRITI)*, IEEE, Dec. 2021, pp. 570–574. doi: 10.1109/ISRITI54043.2021.9702838.

[18] J. Devlin, M.-W. Chang, K. Lee, K. T. Google, and A. I. Language, "BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding." [Online]. Available: https://github.com/tensorflow/tensor2tensor

Advance Sustainable Science, Engineering and Technology (ASSET)
Vol. 6, No.2, April 2024, pp. 0240204-01 ~ 0240204-06
ISSN: 2715-4211 DOI: https://doi.org/10.26877/asset.v6i2.18103

Implementation of Data Layer In Blockchain Network Using SHA256 Hashing Algorithm

Clivent Gerhard Sondakh*, Rizka Ardiansyah, Yuri Yudhaswana Joefrie, Dwi Shinta Angreni, Mohammad Yazdi Pusadan

Information Technology Department, Faculty of Engineering, Tadulako University, Jl. Soekarno Hatta No KM 9 Palu 94148, Central Sulawesi, Indonesia

*clivent1@gmail.com

Abstract. The escalating demand for secure data management in blockchain systems has prompted the exploration of advanced cryptographic techniques. Leveraging the SHA256 hashing algorithm, this implementation aims to fortify data integrity, confidentiality, and authentication within the blockchain network. By meticulously examining the algorithm's application, the research demonstrates its efficacy in ensuring tamper-resistant data storage and retrieval, quantifying improvements in security percentages and specific metrics. The integration of SHA256 within the data layer is explored in technical detail, highlighting the concrete benefits of heightened security and immutability. The analysis discusses practical implications and delves into potential advancements in blockchain technology, offering valuable insights for researchers, developers, and practitioners seeking to bolster the robustness of data layers in blockchain networks.

Keywords: Blockchain, Cryptography, Data Layer, Hashing, Algorithm

(Received 2024-01-08, Accepted 2024-01-22, Available Online by 2024-03-08)

1. Introduction

Blockchain has become a focal point of discussion for its potential to introduce novel technology and offer a new approach to obtaining and sharing information [1]. Functioning as a decentralized and immutable system, it distinguishes itself from traditional systems, incorporating ideas from years of research [2,3,4].

The need for new regulations arises to address current challenges and transform the traditional system into a decentralized and immutable one [5,6]. Recognizing shortcomings in the existing framework, there is a call for the development of a new regulatory structure. Emphasizing the importance of tracking and securing transactions within the blockchain system, the text underscores the need for trust in transactions within a distributed environment [7].

The focus centers on the importance of trust in transactions within a distributed environment, achievable through blockchain technology. Specific references are made to concealing user information through cryptography and utilizing the SHA256 hashing algorithm for security [8].

A crucial element enabling the effective functioning of the blockchain system is its ability to track all individual actions and ensure security. The use of the SHA256 hashing algorithm in blockchain technology plays a pivotal role in maintaining security by assigning a unique code to each piece of information. This code, proven to be highly secure through testing, ensures the effective functioning of the system.

This research aims to implement a Data Layer on the Blockchain Network Using the SHA256 Hashing Algorithm. The results of this research are expected to contribute to knowledge and insights in enhancing the resilience of the data layer on the blockchain network.

2. Methods

2.1. Type of Research

This research employs an exploratory research design to investigate a topic without specific hypotheses, leading to the expansion of insights and the cultivation of a deeper understanding of a less-explored subject. This research serves as a fundamental starting point for future investigations, providing a foundation for broader and more in-depth research initiatives [9-11]. The exploratory research conducted aims to examine the implementation of data layers on the blockchain network using the SHA256 hashing algorithm, thereby enhancing knowledge and insights into improving the resilience of data layers on the blockchain network as a basis for future research.

2.2. System Development

This research employs different phases of the Waterfall Model to guide the development process [12]. This model was chosen because it provides obvious specifications, minimal likelihood of significant changes, and low-performance variability in this study. The structured framework of this model offers clarity in the stages and documentation required to ensure planned and efficient implementation. Starting with a system requirements analysis, the research progresses through design, implementation, and testing, culminating in deployment and maintenance. This sequential approach allows for a thorough examination of each aspect of the system.

3. Results and Discussion

3.1. Planning

The planning phase involves defining the project scope, conducting a comprehensive analysis of requirements, designing a technical architecture, allocating necessary resources, establishing a realistic timeline, assessing, and mitigating risks, planning a thorough testing strategy, developing a deployment plan, creating detailed documentation, and providing training for team members [13]. This strategic planning ensures a systematic and efficient implementation process, minimizing risks and facilitating the successful integration of SHA256 into the blockchain data layer while maintaining the security and integrity of data transactions.

3.2. Analysis

The analysis phase involves a thorough examination of the project requirements and considerations. During this phase, the specific needs, and objectives for incorporating SHA256 are identified, and potential challenges or constraints are analyzed. Stakeholder input is gathered to understand both functional and non-functional requirements, ensuring that the implementation aligns with security standards and performance expectations [14]. This analysis phase serves as the foundation for subsequent planning and technical design, guiding the development team in crafting a detailed and effective strategy for integrating SHA256 into the data layer.

3.3. System Design

In the system design as shown in Figure 1 data will be stored in blocks that are linked as a blockchain, these blocks represent the historical changes of the data and can't be altered or removed, SHA256 hashing algorithm is employed for the storage of the blockchain within the consensus mechanism [15]. This consensus

mechanism involves a hashing process that combines the hash of the previous data, the current data, and a random value [16]. The integration of consensus protocol is a fundamental aspect of many blockchain systems. In this case is a Proof of Work protocol, consensus must find a nonce value that, when hashed with the data of the block, produces a hash that meets certain predefined criteria. The SHA256 algorithm plays a crucial role in this process, providing a deterministic and secure way to generate the block's hash. The difficulty of the mathematical puzzle is adjusted dynamically to ensure that the time required to find a valid nonce aligns with the desired block generation rate. As illustrated in Figure 2, a new block depends on the hash of the preceding block, ensuring that each block has a relationship, making each block interdependent with the others. This ensures that any change in the data will result in a different hash value.

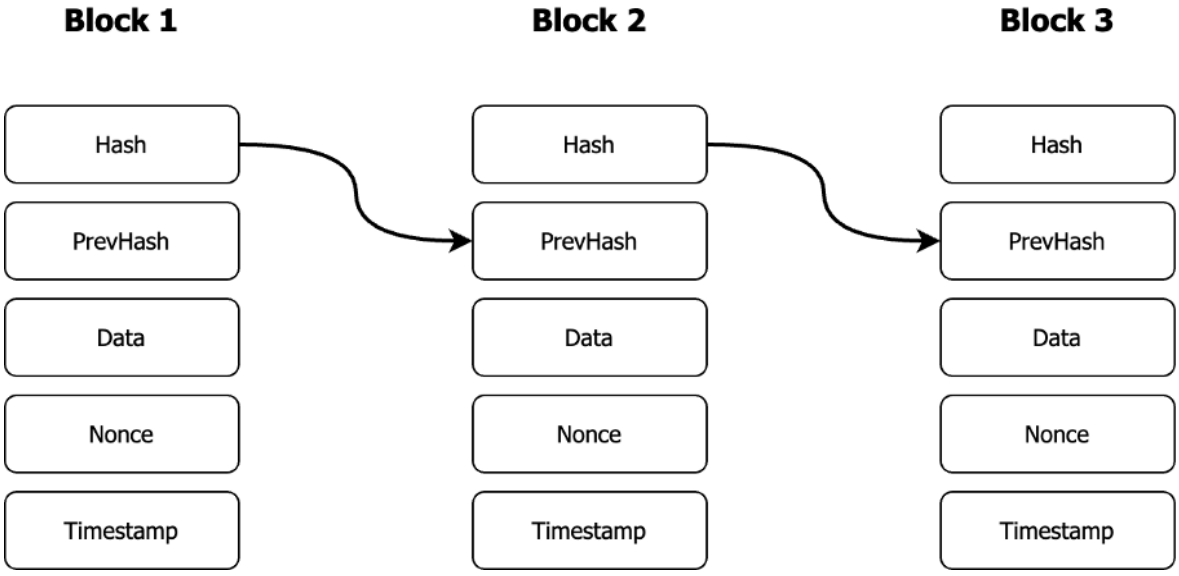


Figure 1. Blockchain Data Layer Design

3.4. Implementation

The implementation of the blockchain's data layer was conducted using Go programming language and using SHA256 algorithm for hashing method. The block structure consists of block hash, the previous block hash (except first/genesis block), data that will be stored inside the block, and timestamp.

Figure 2. Block Structure

The block constructor method will need to satisfy a consensus protocol function interface, this will enable

the function to not have to care what kind of consensus protocol is being used and can be easily swapped if needed.

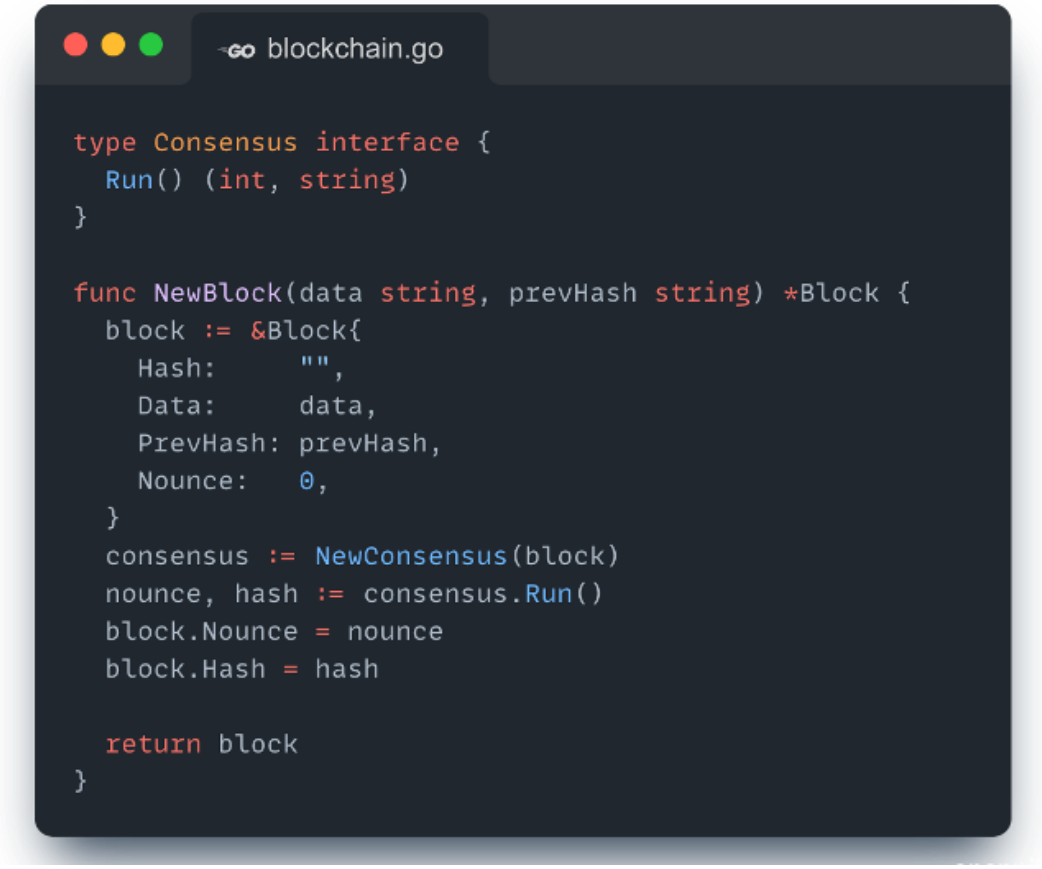


Figure 3. Block Constructor & Consensus Interface

Consensus function which in this case is using Proof of Work consensus protocol, will run the hashing algorithm that will take a Difficulty constant to and data as a parameter, these parameters will determine the hash output [17]. Figure 4 gives the example of block output, changing a single data in this block will result in a different output of the hash, which makes the blockchain invalid since the previous hash that was linked to this block didn't match with this block hash.

Data inside the block will be stored as string type, resulting from JSON serialization; this enables users to store data in their own structure if it is a valid JSON object. This data can later be deserialized for data reading purposes [18].

3.5. System Testing

In the testing phase, this system functions to test the implemented system to ensure it operates as expected and to identify any existing bugs in the system. The testing of this system is conducted using black-box testing. Black-box testing is a type of testing that focuses on the functions of the existing system.

The results of the system testing indicate success in handling block changes, a crucial aspect of system resilience. The interconnected and interdependent nature of blocks in a blockchain system means that any change to one block has the potential to affect others. The system's ability to manage these changes demonstrates a robust design and implementation, ensuring the integrity of the entire blockchain structure. Despite these successes, it is essential to acknowledge potential challenges or limitations that may not be immediately evident from the test results. This may involve considerations such as scalability, transaction volume in the real world, or potential vulnerabilities that may not emerge in a controlled testing environment. Addressing these aspects comprehensively is critical to ensuring system readiness for real-world implementation and continued effectiveness in dynamic operational scenarios.

Table 1. System Testing Results

Test Case	Description	Result
Hash	Hash with the SHA256 Algorithm.	Success
Change Block	When one block changes, other blocks will also be affected because they are interconnected and dependent on each other.	Success

4. Conclusion

Through a meticulous examination of the algorithm's implementation in the data layer, this study has showcased its effectiveness in ensuring tamper-resistant data storage and retrieval. The integration of SHA256 not only contributes to heightened security but also underscores the principle of immutability, a cornerstone of blockchain technology. The unique codes generated by the algorithm serve as digital fingerprints, providing a secure and verifiable means of tracking and securing transactions within the decentralized environment. The findings of this research hold practical implications for the blockchain community, offering valuable insights for researchers, developers, and practitioners seeking to bolster the robustness of data layers in blockchain networks. The implementation of a Data Layer using the SHA256 hashing algorithm establishes a foundation for a more secure and trustworthy blockchain ecosystem.

In essence, the implementation of the Data Layer on the Blockchain Network using the SHA256 Hashing Algorithm marks a significant step towards fortifying the resilience of blockchain data layers, fostering a secure, efficient, and transparent foundation for the future of decentralized systems. This research endeavors to contribute to the collective knowledge and understanding of blockchain technology, guiding its trajectory toward enhanced security and innovation.

Future research recommendations should prioritize scalability solutions, explore privacy-preserving techniques, improve smart contract security, standardize interoperability, and improve user interfaces. These focus areas will contribute to a more robust, secure, and user-friendly decentralized ecosystem, thereby driving broader adoption and innovation.

References

[1] Raharjo, B. (2021). Fintech Financial Technology Digital Banking. Prima Agus Teknik Foundation Publishers, 1-299.

[2] Aini, Q., Lutfiani, N., & Zahran, M. S. (2021). Analysis of ilearning gamification based on blockchain technology. ADI Interdisciplinary Digital Business Journal, 2(June 1), 79-85.

[3] Sunarya, P. A. (2022). Application of Certificates in Security Systems using Blockchain Technology. MENTARI Journal: Management, Education and Information Technology, 1(1), 58-67.

[4] Raharjo, B. (2022). Future Money: Blockchain, Bitcoin, Cryptocurrencies. Prima Agus Teknik Foundation Publishers, 1-68.

[5] Isaac, D. (2022). Review of Education Decentralization Policy in Indonesia. PAPATUNG: Journal of Public Administration, Government and Politics, 5(1), 30-36.

[6] Wahyudi, M. A., & Lutfi, A. (2019). Analysis of educational reform in realizing equal distribution of educational quality in Indonesia. Journal of Public Administration (Public Administration Journal), 9(2), 191-201.

[7] Mahendra, B. A. (2023). Analisis Strategi Pengembangan Teknologi Blockchain Sebagai Media Transparansi Wakaf Di Badan Wakaf Indonesia (Doctoral dissertation, Universitas Islam Sultan Agung).

[8] Putra, I. S. Implementation of the BLAKE3 Hashing Algorithm on JSON Web Token (JWT) for a REST-API Based Web Application Authentication Mechanism.

[9] Setiana, D. S., Ayuningtyas, A. D., Wijayanto, Z., & Kusumaningrum, B. (2021). Exploration of

ethnomathematics at the Yogyakarta Palace Railway Museum and its integration into mathematics learning. Ethnomathematics Journal, 2(1), 1-10.
[10] Muhajirin, M., & Panorama, M. (2017). PRACTICAL APPROACH; Qualitative and Quantitative Research Methods.
[11] Priadana, M. S., & Sunarsi, D. (2021). Quantitative Research Methods. Pascal Books.
[12] Satrio, F. D., & Rismayadi, A. A. (2021). ANDROID BASED PRODUCTIVITY ASSURANCE REPORT APPLICATION AT PT. TELKOM ACCESS. eProceedings of Informatics Engineering (PROTECTIVE), 1(1), 179-185.
[13] Nuralam, I. P., Aini, E. K., Ramadhan, H. M., & Asmoro, P. S. (2023). Introduction to Marketing Research: Uncovering Effective Research Practices. Brawijaya University Press.
[14] Aditya, A. (2023). Software Requirements Engineering. SOFTWARE ENGINEERING: CONCEPTS, METHODS AND BEST PRACTICES.
[15] Liza, W., Thomas, S., & Chrisdityra, L. (2022). Implementation of a proof-of-work consensus algorithm in blockchain against medical records. Pekommas Journal, 7(1), 41-52.
[16] Tahir, C., Airlangga, G., & Hendrawan, K. Bitcoin: A Peer-to-Peer Electronic Cash System.
[17] Sutopo, A. H. (2023). Unlocking the Future: Building Web3 Websites with Unstoppable Domains. Topazart.
[18] Purnomo, R. F., Purbo, O. W., & Aziz, R. A. (2021). Firebase: Building Android-Based Applications. Andi Publisher.

Advance Sustainable Science, Engineering and Technology (ASSET)
Vol. 6, No.2, April 2024, pp. 0240205-01 ~ 0240205-08
ISSN: 2715-4211 DOI: https://doi.org/10.26877/asset.v6i2.18129

Aspect-based Sentiment Analysis on Electric Motorcycles: Users' Perspective

Muchamad Taufiq Anwar[1*], Denny Rianditha Arief Permana[1], Ahmad Juniar[1], Anggy Eka Pratiwi[2]

[1]Automotive Industry Information System, Politeknik STMI Jakarta, Jl. Letjen Suprapto No. 26, Central Jakarta 10510, Jakarta, Indonesia

[2]Department of Computer Science and Engineering, Indian Institute of Technology Jodhpur, Jodhpur, India

*taufiq@stmi.ac.id

Abstract. Electric Vehicles (EVs) adoption is emerging especially electric motorcycles due to their lower price. Research has shown that the majority of people have positive sentiments towards EVs but most of the sentiments were from people who did not already own or use EVs, but rather from people who reacted / commented towards a product that is recently being launched/announced. This research aims to evaluate users' opinions regarding the positive and negative aspects of electric motorcycles they had purchased / used. This information will be beneficial for the manufacturers and marketers as an evaluation for their products; and it is also beneficial for prospective buyers as a buying consideration. This research uses Aspect-Based Sentiment Analysis applied on 844 electric motorcycles review data from www.bikewale.com website. Results showed that the notable positive sentiments are related to smooth riding experience and low maintenance. Whereas notable negative sentiments are related to poor build quality and product malfunctions. The other aspects of electric motorcycles received mixed sentiments such as related to vehicle speed and customer service. The research findings, limitations, and future research direction are discussed.

Keywords: aspect-based sentiment analysis, electric vehicle, electric motorcycle, users' perspective

(Received 2024-01-28, Accepted 2024-02-05, Available Online by 2024-03-08)

1. Introduction

Electric Vehicles (EVs) are emerging and vehicle manufacturers have released their products to the public, more specifically electric cars and electric motorcycles (EMs). Having lower price tags, electric motorcycle ownership has gained a more rapid growth than electric cars. When purchasing electric motorcycles, the public has their own expectation about the products, especially related to their expected / promised

advantages. After the purchase has been made and the product being used, the owner / user will have their real-life experience and opinion about the product. The users will now see the real advantages and disadvantages of the product. This research aims to evaluate users' opinions regarding the positive and negative aspects of electric motorcycles they had purchased / used. This information will be beneficial for both the manufacturer, marketers, and prospective buyers [1]. As for prospective buyers, it will inform them about the actual advantages and disadvantages of the product. As for marketers, it is useful for determining what features should be advertised. Whereas for the manufacturers, it serves as an evaluation of which aspects of the product are already good / well-received by the buyer and which aspects need improvement. Previous research commonly focused on EV buying motivation / considerations [2], and not on the actual advantages and disadvantages of EVs. Therefore this research aims to identify the real-life advantages and disadvantages of EVs as mentioned by people who already bought / used the vehicle. To identify the real-life advantages and disadvantages of EVs, aspect-based sentiment analysis is used as the research approach.

Sentiment Analysis (SA) is one of the Natural Language Processing (NLP) tasks where people's opinions/perceptions (positive or negative) toward certain entities such as products and services are extracted and analyzed [3]–[5]. There are three levels of SA: document-level SA, sentence-level SA, and aspect-based SA (ABSA) [6]. The document-level and sentence-level SA can only classify each document/sentence into a class and is incapable of finding the more fine-grained sentiment for a specific aspect of interest, such as the comfort and the speed of a vehicle. The document-level and sentence-level SA will struggle to classify a document/sentence that contains both positive and negative sentiments such as "The range is good but it's too pricey.". They will miss the details for each positive and negative sentiments. In this case, ABSA has the advantage to identify the aspects of the product that are perceived as positive or negative by the users / customers. One of the data sources that can be used for SA / ABSA is based on review data that is available online [7], [8]. In this research, ABSA is applied to EM review data to extract the positive and negative aspects of the already-owned EM.

The three most commonly used approaches in ABSA are lexicon-based, traditional machine learning, and deep learning approach [6]. This research will use the lexicon-based approach, which uses a predefined dictionary of words and phrases that are associated with a positive or negative sentiment polarity. The ABSA in this research will use a lexicon-based sentiment analysis tool called VADER (Valence Aware Dictionary and sEntiment Reasoner) where the sentiment classification is based on positive or negative keywords such as "great", "good", "worst", and "poor" [9].

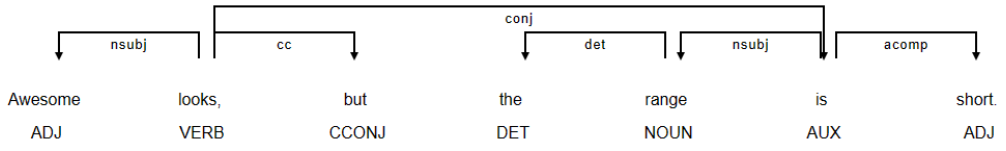
Research has shown that the public generally has positive sentiments towards EV [10], [11]. Sentiment analysis regarding EVs has been carried out based on social media data [1], [10]–[12]. The limitation of such data source is that it cannot guarantee that the person is already using the EVs. For example, research by Anwar (2023) [11] showed that most of the sentiments from social media are those from people who did not already own or use EVs, but rather people's reaction towards a product that is recently being launched/announced [11], [12]. Also, the aspect being discussed is usually unrelated to the aspect of the product itself, but rather on other aspects such as EVs being environmentally friendly [11] and EV price [12]. Other research also showed that multiple users may give their opinions on different aspects of EVs, for example, potential customers talk about the price, brands of EVs, driving range, and charging time [13]. These kinds of sentiment analysis results do not reflect the real-world experience in using the product, which the information is much needed especially for people who are planning to buy an EV. To extract the specific aspects of the product that need paying attention in buying consideration, we need to use Aspect-Based Sentiment Analysis (ABSA) instead of document-level sentiment analysis. Therefore, this research will focus on conducting ABSA on EV reviews from people who already use or own an EV. Other research that used Chinese EV review data had performed sentiment analysis but on document-level [14].

2. Methods

The research methods are shown in Figure 1. The electric motorcycle review data was collected from www.bikewale.com, an Indian website that provides information related to motorcycles including specifications, price, and reviews from the users. In this research, the data being used is the review text. The review data is collected across multiple electric motorcycle brands and models to inform us of general sentiments related to Electric Motorcycles. The text will undergo dependency parsing (identifying grammatical relationships between words in a sentence) and noun phrase extraction using

the SpaCy [15] module in Python. SpaCy is known to be the fastest NLP parser available while and has the highest state-of-the-art accuracy [16]–[18]. The parser used in SpaCy is a transition-based dependency parser. It uses the large English pipeline named "en_core_web_lg" which was trained using three datasets namely the OntoNotes Release 5.0 (University of Pennsylvania), ClearNLP (Emory University), and WordNet 3.0 (Princeton University) [15]. The accuracy of the Part-of-speech tagger component is 0.97, 0.92 for Sentence segmentation precision, and 0.90 for Labeled dependencies [15]. The extracted noun phrases are then classified into positive and negative sentiments using the VADER SentimentIntensityAnalyzer [9] submodule in the nltk [19] module in Python. VADER is chosen because of its rule-based approach which brings four major advantages: transparency, domain-independence, no training requirement, and speed. First, due to its rule-based nature, VADER is transparent and allows us to understand why a particular phrase was assigned to a positive or negative sentiment. Second, the traditional Machine Learning and Deep Learning model requires training on labeled and specific domain datasets, whereas VADER does not require training and is domain-independent which will work on any domain. Lastly, VADER is faster and more lightweight than the Deep Learning model. Finally, the frequency of each noun phrase is calculated to extract the aspect-based sentiment and the result is interpreted. The experiments were carried out using Python 3.7 and Jupyter Notebook [20]. An example of the research process is shown in Table 1.

Table 1. Example of dependency parsing, noun phrase extraction, and sentiment classification

Phase	**Example**
Review text	Awesome looks, but the range is short..
Dependency parsing	
Extracted noun phrases	"awesome looks", "short range"
Sentiment classification	"awesome looks" -> positive sentiment "short range" -> negative sentiment

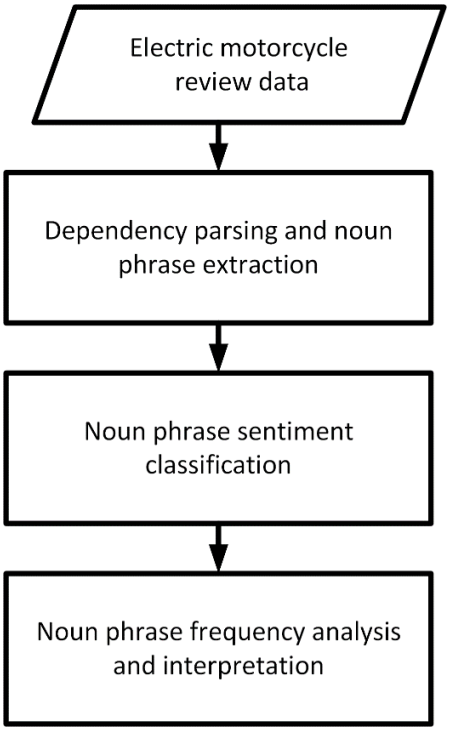


Figure 1. Research Methods

3. Results and Discussion

The data acquisition resulted in a total of 844 rows of review data from multiple EM brands and models with 79% of the reviews being from owners. The rest of them are either people who use other people's EM or just give a comment about a particular EM product. The length of the reviews is up to 5500 characters or 970 words with an average of 462 characters or 89 words. On the document-level sentiment analysis, our data showed that most of the reviews are positive (72%). The results of the positive and negative aspects are shown in Table 2 and Table 3 respectively. The frequency of aspects mentioned in the review can reflect the importance of each aspect which can inform the manufacturer, marketer, or consumer about which aspects need paying attention [21].

Table 2 shows that the notable positive sentiments are related to a smooth riding experience and low maintenance. Research has shown that riding comfort is related to sound and vibration [22]. This is reasonable since EV uses an electric motor which produces minimal sound and vibration. The result related to EV's low maintenance is in line with other research results which mentioned that one of the advantages of EVs is low maintenance [23]. EVs are also known to have lower maintenance costs compared to internal combustion engine (ICE) vehicles, although it has higher initial costs [24]. The low maintenance of EVs comes from the use of battery and electric motors [25] instead of the more complicated drivetrain system in combustion-based vehicles.

On the other hand, the notable negative sentiments are related to poor build quality and product malfunction. The low build quality may come from the competition of the manufacturer to offer lower prices since EVs are generally more expensive due to battery cost [26]. Even though many people mentioned the poor build quality, there are quite a lot of people who mentioned that the build quality is good. This can be explained by the fact that they reviewed different brands and models. Related to EV malfunction, it can be mitigated by using a Fault Detection system especially if it is related to the battery and electric motor as they are the most critical components [27]. Other aspects such as customer service

and EM speed received mixed sentiments. This variation may also be attributed to people reviewing different brands/models.

Contrary to other research results, our result does not mention the sentiment towards charging stations [28]. It might be explained by that EM can be charged at home and is used for shorter travel distances than cars, thus it does not need public charging stations. Charging EM at home is generally much easier to do, much less costly to install, and does not need higher power than an EV car charger.

Other research in EV (car) sentiments analysis using LDA found that the most important topics/aspects the customers consider when planning to purchase an EV are: dynamics (top speed, acceleration, etc), technology (driving range, charge time, etc), safety (braking properties, operating stability, etc), comfort (suspension, seat, etc), and cost [29]. The result of our research agreed on four out of the five aspects, excluding safety. The lack of findings about the safety aspect in our research aspect might be explained by that the EM owner is less concerned about the vehicle safety relative to other aspects of the vehicle, or that the safety of electric motorcycles is less of a concern than those of electric cars. Nevertheless, future research may need to explore the safety aspect of EM as it is an important aspect to consider, especially regarding government policy related to vehicle safety. Other ABSA research on ICE cars use has found that most of its users are satisfied with several positive aspects of the vehicle such as driving comfort, good fuel economy / mileage, reliability, value for money, quiet ride, good acceleration, well-designed, solid build, etc [30]. Many of these aspects have also been captured in our research. Manufacturers may use this information to target the specific aspect they need to maintain / improve the quality of.

Compared to the previous research [29], this research found three new important aspects related to EVs: build quality, product malfunction, and customer service. All of these aspects have mixed sentiments but generally tend to be on the negative side. All of this negativity might be attributed to price competition among manufacturers/companies. Therefore, based on this research result, the manufacturer/company needs to pay attention to product build quality and electric system reliability; whereas marketers/dealers should focus on giving good service to their customers.

The limitation of this research is that this research is based on short-term reviews where most of the reviews (70%) are from owners who give reviews within 1 year of EM usage/ownership. Future research may need to perform follow-up research with a longer period of ownership, such as 3-5 years of use, especially regarding the battery performance because the battery capacity might be degraded after 3-5 years of use. Some of the electric motorcycle parts may also degrade and need replacement after a certain period of usage time. This information is also important to consider for people planning to buy an EV.

Table 2. List of positive aspects and its corresponding frequency

Positive aspects	Frequency
Good/smooth riding experience	22
Good service	13
Low maintenance	11
High speed	11
Good quality	10
Good mileage	8
Amazing look	4
Large boot space	3
Functional buttons	2
Affordable price	2
Removable battery option	2

Table 3. List of negative aspects and its corresponding frequency

Negative aspects	Frequency
Poor build quality	57
Bad experience	15
Many problems	11
Bad customer service	11
Low speed	6
Bad salesperson	5
Bad servicing	4
Suddenly stops	3
Frequent price hikes	2

4. Conclusion

This research aims to uncover users' opinions on the positive and negative aspects of electric motorcycle usage based on user reviews. Results showed that the notable positive sentiments are related to a smooth riding experience and low maintenance. Whereas notable negative sentiments are related to poor build quality and product malfunctions. The other aspects of EM received mixed sentiments such as related to vehicle speed and customer service. Based on the results, manufacturers should improve the overall build quality of their products and the quality of their electrical systems. As for the marketers, they should provide better customer service in handling prospective buyers, handling customer complaints, and giving after-sales services. As for the prospective buyers, they should carefully choose which brand / model has a good build quality and good customer service. The limitation of the study is that this research uses short-term reviews in which most of the reviews (70%) are from owners who give reviews within 1 year of using/owning the electric motorcycles. Future research may need to perform a follow-up research on a longer ownership period, such as 3-5 years of use, as people's opinions may change over time and some of the components especially the battery may degrade in performance and become a significant issue. Future research may focus on exploring some aspects related to longer-term reviews such as battery degradation, vehicle performance degradation, and component degradation.

References

[1] R. Jena, "An empirical case study on Indian consumers' sentiment towards electric vehicles: A big data analytics approach," *Ind. Mark. Manag.*, vol. 90, pp. 605–616, 2020.

[2] M. T. Anwar, M. P. Utami, L. Ambarwati, and A. W. Arohman, "Identifying Social Media Conversation Topics Regarding Electric Vehicles in Indonesia Using Latent Dirichlet Allocation," in *2022 IEEE International Conference on Cybernetics and Computational Intelligence (CyberneticsCom)*, 2022, pp. 102–106.

[3] M. Birjali, M. Kasri, and A. Beni-Hssane, "A comprehensive survey on sentiment analysis: Approaches, challenges and trends," *Knowledge-Based Syst.*, vol. 226, p. 107134, 2021.

[4] I. Chaturvedi, E. Cambria, R. E. Welsch, and F. Herrera, "Distinguishing between facts and opinions for sentiment analysis: Survey and challenges," *Inf. Fusion*, vol. 44, pp. 65–77, 2018.

[5] M. T. Anwar, L. Ambarwati, D. Agustin, and others, "Analyzing Public Opinion Based on Emotion Labeling Using Transformers," in *2021 2nd International Conference on Innovative and Creative Information Technology (ICITech)*, 2021, pp. 74–78.

[6] H. Liu, I. Chatterjee, M. Zhou, X. S. Lu, and A. Abusorrah, "Aspect-based sentiment analysis: A survey of deep learning methods," *IEEE Trans. Comput. Soc. Syst.*, vol. 7, no. 6, pp. 1358–1375, 2020.

[7] Y. Yiran and S. Srivastava, "Aspect-based Sentiment Analysis on mobile phone reviews with LDA," in *Proceedings of the 2019 4th International Conference on Machine Learning*

Technologies, 2019, pp. 101–105.
[8] Y. Zhang, J. Du, X. Ma, H. Wen, and G. Fortino, "Aspect-based sentiment analysis for user reviews," *Cognit. Comput.*, vol. 13, no. 5, pp. 1114–1127, 2021.
[9] C. Hutto and E. Gilbert, "Vader: A parsimonious rule-based model for sentiment analysis of social media text," in *Proceedings of the international AAAI conference on web and social media*, 2014, vol. 8, no. 1, pp. 216–225.
[10] F. J. Costello and K. C. Lee, "Exploring the Sentiment Analysis of Electric Vehicles Social Media Data by Using Feature Selection Methods.," *J. Digit. Converg.*, vol. 18, no. 2, 2020.
[11] M. T. Anwar, "Analisis Sentimen Masyarakat Indonesia Terhadap Produk Kendaraan Listrik Menggunakan VADER," *JATISI (Jurnal Tek. Inform. dan Sist. Informasi)*, vol. 10, no. 1, pp. 783–792, 2023.
[12] D. P. Demirer and A. Büyükeke, "Analysing perceptions towards electric cars using text mining and sentiment analysis: A case study of the newly introduced TOGG in Turkey," *Appl. Mark. Anal.*, vol. 7, no. 4, pp. 386–399, 2022.
[13] T. Ruan and Q. Lv, "Public perception of electric vehicles on reddit over the past decade," *Commun. Transp. Res.*, vol. 2, p. 100070, 2022.
[14] M. Wang, H. You, H. Ma, X. Sun, and Z. Wang, "Sentiment Analysis of Online New Energy Vehicle Reviews," *Appl. Sci.*, vol. 13, no. 14, p. 8176, 2023.
[15] H. Matthew, I. Montani, S. Van Landeghem, and A. Boyd, "spaCy: Industrial-strength Natural Language Processing in Python," 2020, doi: 10.5281/zenodo.1212303.
[16] N. Colic, "Dependency parsing for relation extraction in biomedical literature," *Master's thesis, Univ. Zurich, Switz.*, 2016.
[17] N. Colic and F. Rinaldi, "Improving spaCy dependency annotation and PoS tagging web service using independent NER services," *Genomics \& informatics*, vol. 17, no. 2, 2019.
[18] M. Honnibal and M. Johnson, "An improved non-monotonic transition system for dependency parsing," in *Proceedings of the 2015 conference on empirical methods in natural language processing*, 2015, pp. 1373–1378.
[19] S. Bird, E. Klein, and E. Loper, *Natural language processing with Python: analyzing text with the natural language toolkit*. " O'Reilly Media, Inc.," 2009.
[20] T. Kluyver *et al.*, "Jupyter Notebooks -- a publishing format for reproducible computational workflows," in *Positioning and Power in Academic Publishing: Players, Agents and Agendas*, 2016, pp. 87–90.
[21] B. Gong, R. Liu, X. Zhang, C.-T. Chang, and Z. Liu, "Sentiment analysis of online reviews for electric vehicles using the SMAA-2 method and interval type-2 fuzzy sets," *Appl. Soft Comput.*, vol. 147, p. 110745, 2023.
[22] X. Wang, A.-L. Osvalder, and P. Höstmad, "Influence of sound and vibration on perceived overall ride comfort—A comparison between an electric vehicle and a combustion engine vehicle," *SAE Int. J. Veh. Dyn. Stability, NVH*, vol. 7, no. 10-07-02–0010, pp. 153–171, 2023.
[23] A. Thattil, S. Vachhani, D. Raval, P. Patel, and P. Sharma, "Comparative study of using different electric motors for EV," *Int. Res. J. Eng. Technol.*, vol. 6, no. 4, pp. 4601–4604, 2019.
[24] I. Veza, M. Z. Asy'ari, M. Idris, V. Epin, I. M. R. Fattah, and M. Spraggon, "Electric vehicle (EV) and driving towards sustainability: Comparison between EV, HEV, PHEV, and ICE vehicles to achieve net zero emissions by 2050 from EV," *Alexandria Eng. J.*, vol. 82, pp. 459–467, 2023.
[25] K. A. Nitesh and Ravichandra, "A study on battery controller design for the estimation of state of charge (SoC) in battery management system for electric vehicle (EV)/hybrid EV (HEV)," *SN Comput. Sci.*, vol. 2, no. 3, p. 197, 2021.
[26] A. König, L. Nicoletti, D. Schröder, S. Wolff, A. Waclaw, and M. Lienkamp, "An overview of parameter and cost for battery electric vehicles," *World Electr. Veh. J.*, vol. 12, no. 1, p. 21, 2021.
[27] M. Z. Khaneghah, M. Alzayed, and H. Chaoui, "Fault detection and diagnosis of the electric

motor drive and battery system of electric vehicles," *Machines*, vol. 11, no. 7, p. 713, 2023.
[28] O. I. Asensio, K. Alvarez, A. Dror, E. Wenzel, and C. Hollauer, "Evaluating popular sentiment of electric vehicle owners in the United States with real-time data from mobile platforms."
[29] X. Ren, S. Sun, and R. Yuan, "A study on selection strategies for battery electric vehicles based on sentiments, analysis, and the MCDM model," *Math. Probl. Eng.*, vol. 2021, pp. 1–23, 2021.
[30] M. T. Anwar, D. Trisanto, A. Juniar, and F. A. Sase, "Aspect-based Sentiment Analysis on Car Reviews Using SpaCy Dependency Parsing and VADER," *Adv. Sustain. Sci. Eng. Technol.*, vol. 5, no. 1, p. 230109, 2023.

Advance Sustainable Science, Engineering and Technology (ASSET)
Vol. 6, No.2, April 2024, pp. 0240206-01 ~ 0240206-09
ISSN: 2715-4211 DOI: https://doi.org/10.26877/asset.v6i2.18309

Identification Of Chatbot Usage In Online Store Services Using Natural Language Processing Methods

Anshar Daud[1*], Dola Irwanto[2], Muh Said[1], Mutiarini Mubyl[1], Mustamin[1]

[1]Instititut Teknologi dan Bisnis Nobel Indonesia
[2]Universitas Pamulang, Banten, Indonesia

*anshar@nobel.ac.id[1]

Abstract. Chatbot is one of the implementations of artificial intelligence in helping human tasks. The way the chatbot itself works is to answer questions directly according to the database that has been created. Chatbot helps online store owners answer questions from the same customers so that there is efficiency in terms of employee salaries. 40 questions that are often asked by customers to the online store admin. In its use, chatbot is one of the implementations of machine learning, where the function of machine learning itself is to improve the ability of machines to learn new information from data and develop the ability of machines to solve problems. Machine learning requires instruction from training data or input to teach machines how to solve problems, answer questions, and draw conclusions from the results of data processing. Furthermore, it is processed using Natural Language Processing (NLP). Natural Language Processing (NLP) is a branch of artificial intelligence that is able to study communication between humans and computers through natural language. The processing stages are identifying the intent, processing the input and displaying the results according to the input. Followed by testing the accuracy level. Then conduct testing using 40 question and answer data. Then obtained 36 answers that are appropriate and 4 answers that are not appropriate with the percentage of accuracy of the answers generated from the chatbot is 90 percent. The results of this test can answer the questions asked by customers. This chatbot can make it easier for customers to get information with a very good level of accuracy.

Keywords: Chattbot, Natural Languange Processing, Artificial Intellegent, Online Store, Service, Machine Learning.

(Received 2024-01-31, Accepted 2024-02-10, Available Online by 2024-03-08)

1. Introduction

Online stores are one of the shifting trends after internet technology became increasingly popular. The shopping community does not need to leave the place, they can choose the items to be purchased through online stores through the marketplace by looking at the photos displayed along with the description of the goods written, Because of this, it has led to an increasing number of online stores in Indonesia, and made online store owners unable to take a lot of profit because if their goods are too expensive buyers will choose to shop at other online stores[1]. Therefore, the owner will try to minimize the costs incurred, one of which is the cost of employee salaries.

Online stores do not require face-to-face meetings with buyers but the meeting is done online using the chat feature. Chatbot is a solution needed by online shop entrepreneurs in order to reduce employee salary costs, especially admin.[2] Chatbot can eliminate the role of admin by replying to questions from customers automatically, so that customers can still be served without having to be done by humans[3].

The development of information technology has brought significant impacts in various aspects of life, including the business world.[4] In the context of e-commerce or online stores, the interaction between customers and sales platforms is crucial to creating a satisfying and efficient experience. In an effort to meet these demands, the application of artificial intelligence technology, particularly chatbots, has become an increasingly popular trend.[5]

A chatbot is a computer system designed to perform human-machine interaction through natural language. [6]Basically, chatbots aim to provide instant responses to user questions or requests without involving human intervention. In the context of online stores, the use of chatbots is not only considered an innovative element, but also a strategy that can improve the efficiency and quality of customer service. One of the growing methods to improve chatbot performance is Natural Language Processing (NLP). NLP is a branch of artificial intelligence that allows machines to understand, interpret, and respond to human language in a manner similar to human-to-human communication. By applying NLP methods to chatbots, it is expected to improve the chatbot's ability to understand and respond to various natural language expressions used by users.[7]

Research on this chatbot gets an accuracy of 90.9 %, while the responsiveness time to answer questions of less than 5 words is 0.01 seconds and 0.02 seconds for more than 5 words.[8] for more than 5 words is 0.02 seconds with a data set of 1000 lines. data set of 1000 lines. This chatbot application is able to answer the questions asked by user, according to the knowledge that has been given previously. The response given by the chatbot depends on the keywords inputted, chatbot will reply to keywords with the most similar pattern from the textual database from the textual database.[9]

Chatbot technology is one form of application with Natural Language Processing (NLP), NLP itself is one of the fields of science. Artificial Intelligence that studying communication between humans and computers through natural language[10]. This application application is known as an automatic conversation agent that running on computer programming or some kind of Artificial Intelligence (AI) interaction between users and machine with the intervention of Natural Language Processing.[11] Chatbots can potentially be called the most promising form of of human-machine interaction that is the most promising and advanced.[12] NLP has many purposes that can assist human communication, such as translation machine and assisting human machine communication, such as with conversational agents and others. A chatbot is a technology whose main purpose is to interact with human users by processing natural language input and generating relative output through a machine-driven rule-driven machine or artificial intelligence engine. Natural language processing uses the stages of tokenizing, filtering, and analysis as well as applying the knuth morris prrat algorithm.[13]

This research aims to identify the use of chatbots in online store services with an NLP approach. Through a deeper understanding of the implementation of this technology, it is expected to reveal the potential for improving service quality, operational efficiency, and customer experience in the context of online stores. Thus, this research not only contributes to the literature on chatbot applications in e-commerce, but also provides practical insights for businesses in utilizing this technology to improve competitiveness and customer satisfaction[14].

2. Methods

The framework in research is an interconnected stage with a systematic arrangement to solve a problem. systematically to solve a problem. [15]At This research framework is carried out in stages to make it easier to solve problems related to related to Chatbot Identification in Improving Online Services Using the Natural Language Processing Method. Work framework in stages of research can be seen in Figure 1[16].

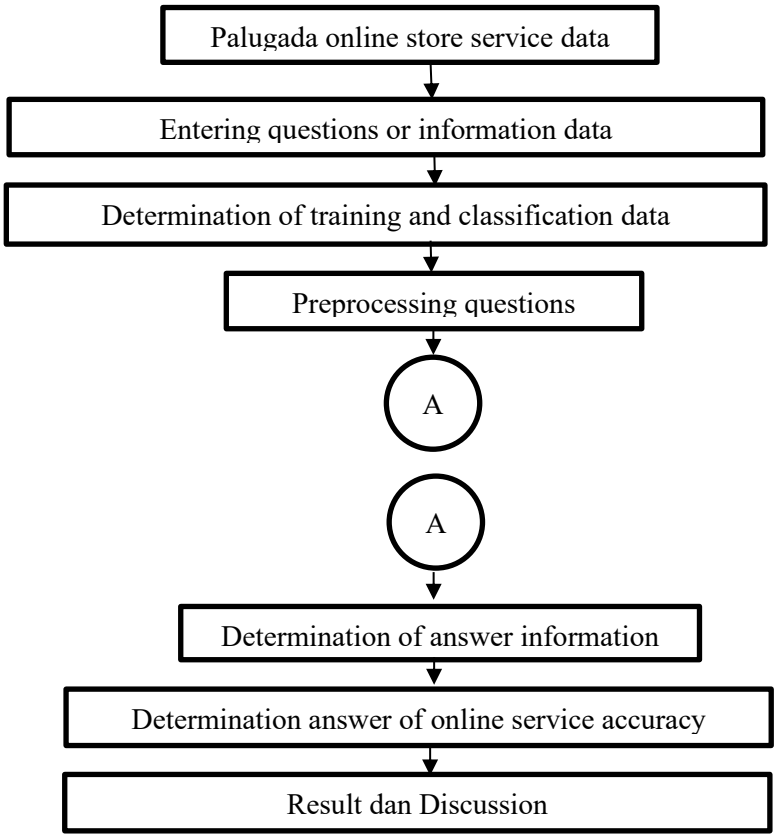


Figure 1. Framework research

2.1 Palugada online store service information data

collection sales information used by the author in this study are:
a. Interviews were conducted with customers and resellers of the palugada online store.
b. Literature studies are carried out by searching for journal references.
c. Questionnaires addressed to customers and resellers of palugada online stores.

2.2 Entering questions or information data
Questions about palugada online shop services are collected along with the answers using the program that has been created. For text-based chatbots, you can type questions or information into the chat or text interface, so that the chatbot will generate a response based on their understanding of your question or information.[17]

2.3 Determination of training and classification data
Training and Classification is done so that the chatbot has the knowledge to understand natural human language, this stage is used when there is one word or keyword that is suspected to be similar, an answer will appear according to customer expectations.[2]

2.4 Preprocessing questions
The preprocessing of these questions is done using algorithms in order to produce appropriate answers in the database. Syntax Analysis in NLP is divided into 3 (three).[8]

a. Case Folding:
   To simplify processing, case folding is often performed to convert all letters in the text to lowercase or uppercase. This helps overcome the issue of case distinction which may not be relevant.
b. Tokenization:
   Tokenization is the process of converting text into smaller pieces, called tokens. Tokenization is an important step in text processing in Natural Language Processing (NLP) by dividing text or sentences into smaller units called tokens. Tokens can be words, phrases, or characters.
c. Filtering or Stopword Removal:
   Filtering or stopword removal is a common step in text pre-processing in Natural Language Processing (NLP). Stopwords are common words that appear frequently in text and tend not to contribute significantly to the meaning of a sentence or document. Stopwords (common words like "the", "and", "is") are often removed to increase the focus on more informative words.

2.5 Determination of answer information.

Chatbot is intended as a tool to answer questions automatically, so before the question is answered, the question and answer must be determined first. Determination of answers in this online store service is collected first, as seen in Figure 2, so that if there is a chatbot question will answer according to the answers that have been provided.

Figure 2. Determination of answer

2.6 Determination answer of online service accuracy.

Determining accuracy based on questions on a chatbot refers to evaluating the performance of the chatbot in answering questions asked by users.[18] Accuracy is a metric that measures the extent to which the answer provided by the chatbot is as expected or desired.[19]

No	Question	Answer
1	Kak, jual apa aja ?	Match
2	Barang, Asli atau KW ?	Match
3	Produknya produksi sendiri?	Match
4	Warna Jilbabnya apa aja ?	Match
5	Harga pas kak ?	Match
6	Ada Ukuran lain ?	Match
7	Apakah Ada garansi ?	Match
8	Bisa retur ?	Match
9	Apakah Gratis Ongkir	Match
10	Apakah ada cashback ?	Match
11	Kalau barang sobek bisa di tukar ?	Match
12	Pengirimannya berapa lama ?	Match
13	Barang saya sampai mana ?	Not match
14	Beli banyak dapat diskon ?	Match
15	Ready Stok ?	Match
16	Ada toko offlinenya ?	Match
17	Cara jadi reseller gimana ?	Match
18	Bahan Jilbabnya apa ?	Match
19	Sama Azahra kualitasnya sama ?	Match
20	Bagaimana jika produknya tidak sesuai dengan gambar ?	Not match
21	Kok mahal ?	Match
22	Bedanya sama Azahra apa ?	Match
23	Ada ukuran XXL untuk bajunya ?	Match
24	Jual baju cowok ?	Match
25	Harga partai minimal pembelian berapa ?	Match
26	Barangku sampai mana kak ?	Match
27	Bagaimana caranya pesan barang ?	Match
28	Lokasi pengiriman dari mana ?	Match
29	Berapa total semuanya ?	Match
30	Boleh aku datang ke toko langsung ?	Match
31	Baju ukuran big size ada ?	Match
32	Apakah ini real pict ?	Not match
33	Bahan baju manohara ini apa ?	Match
34	Ongkir ke surabaya berapa ?	Match
35	Gratis ongkir kak ?	Match
36	Bisa kirim hari ini?	Match
37	Cara cek stok gimana kak ?	Not match
38	Mau jadi reseller syaratnya apa ?	Match
39	Pembayaran COD bisa ?	Match
40	Ini yakin amanah kalau saya transfer dulu ?	Match

Figure 3. Determination answer of online service accuracy.

Accuracy : $\frac{\text{Match answer}}{\text{All answer}} \; x \; 100\%$

: $\frac{36}{40} \; x \; 100\%$

: 90%

Based on testing the accuracy of answers using chattbot, it has been tried by entering 40 questions on the palugada online store service, found 4 answers that do not match so that the accuracy of the chattbot is 90%.

3. Result and Discussion

At this stage the author identifies the chatbot created on WhatsApp to improve online services. The chatbot application is presented in the figure 4.[20]

Figure 4. Result of chatbot

The use of chattbot using the natural language processing method can help the owner in answering questions from customers. as shown in Figure 7, the questions "kak, jual apa aja ?" and "kak jual apa aja ?" the answer will be the same because it uses the natural language processing method to remove the "," sign in the question so that the answer that appears is the same.

3.1 Perform question preprocessing

a. Case Folding is a process in NLP to convert all words into lowercase letters.

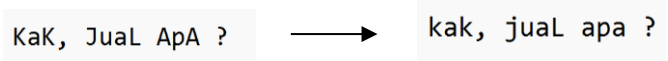


Figure 5. Case Folding Process

Case folding is one of the stages in text processing in Natural Language Processing (NLP). Case folding is used to convert all characters in the text into lowercase or uppercase letters, thus treating the text without considering the difference between uppercase and lowercase letters. This helps to reduce complexity in the next stage of analysis, as the same word with different uppercase or lowercase letters will be considered the same. In Figure 5, it can be seen that there is an uppercase

letter located at the end of the word so that after NLP is applied the uppercase letter that was originally located at the end of the word becomes lowercase. [20]

b. Tokenization which is a process in NLP to break down the input text in the form of questions into words.

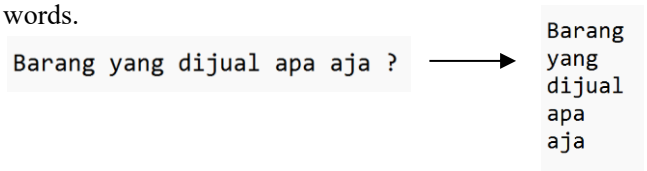


Figure 6. Tokenization Process

Tokenization is the process of dividing text or sentences into smaller parts, called tokens. Tokens can be words, phrases, or other subsections of the text. Tokenization helps transform text into a more structured representation, making analysis and modeling easier.[9] In Figure 6 the sentence "Barang yang dijual apa aja ?" in the Tokenization process will divide the text into tokens in the form of words or collections of words, namely: ["Barang", "yang", "dijual", "apa", "aja"].

c. Filtering, which is a process in NLP to eliminate conjunctions such as "dan"

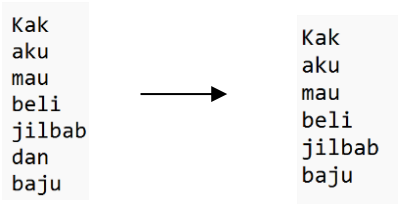


Figure 7. Filtering Process

Filtering in Natural Language Processing (NLP) is the process of removing or filtering certain elements of the text that are considered irrelevant or unwanted. [14]The process carried out at this filtering stage is Stopword Removal where the meaning of stopword removal is common words that often appear in text and tend not to provide significant information. In Figure 7 removes the stopword "and".

4. Conclusion

This research identifies the use of chatbots in online store services by applying Natural Language Processing (NLP) methods. Through this exploration, it can be concluded that the use of chatbots with NLP brings significant positive impacts on various operational aspects and service quality of online stores. The implementation of a chatbot improves operational efficiency by providing instant responses to customer queries and handling routine tasks automatically. This not only reduces the workload of the human customer support team but also speeds up the overall service process. The chatbot's ability to understand natural language using NLP methods provides more natural and relevant interactions with customers. Improvements in understanding synonyms, complex sentence structures, and context enrich the customer experience and result in more accurate responses.

5. References

[1] M. T. Anwar, A. Nurwanda, F. Rahmat, M. Aufal, H. D. Purnomo, and A. Supriyanto, “Fast and Accurate Indonesian QnA Chatbot Using Bag-of-Words and Deep-Learning For Car Repair Shop Customer Service,” *Advance Sustainable Science, Engineering and Technology*, vol. 5, no. 2, p. 0230201, Jul. 2023, doi: 10.26877/ASSET.V5I2.14891.

[2] K. U. Syaliman, N. Fadilah Najwa, T. Informatika, P. Caltex Riau, and S. Informasi, "Chatbot Sebagai Wadah Informasi Perkembangan Covid-19 di Kota Pekanbaru Menggunakan Platform Whatsapp," *IT Journal Research and Development*, vol. 5, no. 2, pp. 234–245, Jan. 2021, doi: 10.25299/ITJRD.2021.VOL5(2).5842.

[3] A. S. Dewantara and J. Aryanto, "Implementation Of A Web-Based Chatbot Using Machine Learning For Question And Answer Services In Universities," *Advance Sustainable Science, Engineering and Technology*, vol. 6, no. 1, p. 0240106, Dec. 2024, doi: 10.26877/ASSET.V6I1.17590.

[4] K. H. TEO and M. J. A. KHIRI, "FCSIT WhatsApp Chatbot," *Trends in Undergraduate Research*, vol. 4, no. 1, pp. c41-51, Jun. 2021, doi: 10.33736/TUR.2866.2021.

[5] M. A. Hakim, "Pembangunan Aplikasi Chatbot Midwify Sebagai Media Pendukung Pembelajaran Ilmu Kebidanan Berbasis Android Di Stikes Bhakti Kencana Bandung," Mar. 2019, Accessed: Dec. 28, 2023. [Online]. Available: http://elibrary.unikom.ac.id

[6] F. Halimatus Sahdiah, E. Fitriani Komara, F. Ekonomi dan Bisnis, and U. Jenderal Achmad Yani, "Penerapan Teknologi Chatbot Whatsapp untuk Meningkatkan Pelayanan dan Jaringan Bisnis," *Jurnal Pemberdayaan Ekonomi*, vol. 1, no. 2, pp. 101–109, Aug. 2022, doi: 10.35912/JPE.V1I2.1234.

[7] M. Mashud and W. Wisda, "Aplikasi Chatbot Berbasis Website sebagai Virtual Personal Assistant dalam Pemasaran Properti," *Inspiration: Jurnal Teknologi Informasi dan Komunikasi*, vol. 9, no. 2, p. 99, Dec. 2019, doi: 10.35585/INSPIR.V9I2.2497.

[8] V. Jonatan, A.-A. Igor, and S. Marcos Lima, "Creation Of A ChatBot Based On Natural Language Proccesing For Whatsapp," *An International Journal (ELELIJ*, vol. 12, no. 2, Oct. 2023, doi: 10.14810/elelij.2023.12402.

[9] M. Mashud and W. Wisda, "Aplikasi Chatbot Berbasis Website sebagai Virtual Personal Assistant dalam Pemasaran Properti," *Inspiration: Jurnal Teknologi Informasi dan Komunikasi*, vol. 9, no. 2, p. 99, Dec. 2019, doi: 10.35585/INSPIR.V9I2.2497.

[10] B. Setiaji, E. Utami, and H. Al Fatta, "Membangun Chatbot Berbasis AIML dengan Arsitektur Pengetahuan Modular," *SEMNASTEKNOMEDIA ONLINE*, vol. 1, no. 1, pp. 18–15, Jan. 2013, Accessed: Dec. 28, 2023. [Online]. Available: https://ojs.amikom.ac.id/index.php/semnasteknomedia/article/view/743

[11] M. Muliyono and S. Sumijan, "Identifikasi Chatbot dalam Meningkatkan Pelayanan Online Menggunakan Metode Natural Language Processing," *Jurnal Informatika Ekonomi Bisnis*, pp. 142–147, Sep. 2021, doi: 10.37034/INFEB.V3I4.102.

[12] E. Larasati Amalia and D. Wahyu Wibowo, "Rancang Bangun Chatbot Untuk Meningkatkan Performa Bisnis," *Jurnal Ilmiah Teknologi Informasi Asia*, vol. 13, no. 2, 2019.

[13] R. N. Astuti and D. M. Fatchan, "PERANCANGAN APLIKASI TEKNOLOGI CHATBOT UNTUK INDUSTRI KOMERSIAL 4.0," *e-Prosiding SNasTekS*, vol. 1, no. 1, pp. 339–348, 2019, Accessed: Dec. 28, 2023. [Online]. Available: https://journal.unusida.ac.id/index.php/snts/article/view/103

[14] M. N. Fauzy and K. Kusrini, "CHATBOT MENGGUNAKAN METODE FUZZY STRING MATCHING SEBAGAI VIRTUAL ASSISTANT PADA PUSAT LAYANAN INFORMASI AKADEMIK," *Jurnal Informa : Jurnal Penelitian dan Pengabdian Masyarakat*, vol. 5, no. 1, pp. 61–67, Aug. 2019, doi: 10.46808/INFORMA.V5I1.70.

[15] Z. M. HANIF, "Pengembangan Aplikasi Whatsapp Chatbot Untuk Pelayanan Akademik Di Perguruan Tinggi," Dec. 2021, Accessed: Feb. 09, 2024. [Online]. Available: https://dspace.uii.ac.id/handle/123456789/37445

[16] Z. Munawar, H. Soerjono, N. I. Putri, Hernawati, and A. Dwijayanti, "Manfaat Kecerdasan Buatan ChatGPT Untuk Membantu Penulisan Ilmiah," *TEMATIK*, vol. 10, no. 1, pp. 54–60, Jun. 2023, doi: 10.38204/TEMATIK.V10I1.1291.

[17] L. Rohmawati, M. A. Nugroho, and W. Wagito, "IMPLEMENTASI CHATBOT PADA WHATSAPP UNTUK MONITORING SUMBER DAYA SERVER," *Journal of Information*

System Management (JOISM), vol. 4, no. 2, pp. 107–112, Jan. 2023, doi: 10.24076/JOISM.2023V4I2.960.

[18] A. Lubis and I. Sumartono, "Implementasi Layanan Akademik Berbasis Chatbot untuk Meningkatkan Interaksi Mahasiswa," *Resolusi : Rekayasa Teknik Informatika dan Informasi*, vol. 3, no. 5, pp. 397–403, May 2023, doi: 10.30865/RESOLUSI.V3I5.767.

[19] K. S. Adi, "Pengembangan Inovasi Layanan Sirkulasi Perpustakaan Berbasis Bot Whatsapp Di Perpustakaan Daerah Kabupaten Purworejo," 2023, Accessed: Feb. 09, 2024. [Online]. Available: https://digilib.uns.ac.id/dokumen/103733/Pengembangan-Inovasi-Layanan-Sirkulasi-Perpustakaan-Berbasis-Bot-Whatsapp-Di-Perpustakaan-Daerah-Kabupaten-Purworejo

[20] M. T. Anwar, A. Nurwanda, F. Rahmat, M. Aufal, H. D. Purnomo, and A. Supriyanto, "Fast and Accurate Indonesian QnA Chatbot Using Bag-of-Words and Deep-Learning For Car Repair Shop Customer Service," *Advance Sustainable Science, Engineering and Technology*, vol. 5, no. 2, p. 0230201, Jul. 2023, doi: 10.26877/ASSET.V5I2.14891.

Advance Sustainable Science, Engineering and Technology (ASSET)
Vol. 6, No.2, April 2024, pp. 0240207-01 ~ 0240207-08
ISSN: 2715-4211 DOI: https://doi.org/10.26877/asset.v6i2.18342

Implementation of the Adaboost Method to Increase the Accuracy of Early Diabetes Predictions to Prevent Death Decision Tree-Based

Laskar Alam*, Ahmad Zainul Fanani, Affandy

Faculty of Computer Science, University of Dian Nuswantoro, Imam Bonjol No. 207 Semarang, Central Java, 50131, Indonesia

*p31202102446@mhs.dinus.ac.id

Abstract. This research discusses the importance of early diabetes prediction and efforts to increase prediction accuracy using a Decision Tree Learning Algorithm and integration of the Adaboost Method. This study uses a data set from Kaggle with 520 records, 16 attributes, and one positive or negative diabetes class. The evaluation method used is the Confusion Matrix. The research results showed that the Decision Tree algorithm achieved an accuracy of 94.23%, but after integrating the Adaboost Method, the accuracy increased to 97.31%. The implications of these findings emphasize the importance of predictive approaches in early disease detection and highlight the potential of the Adaboost method in improving the accuracy of diabetes prediction.

Keywords: Decision Tree, Diabetes, Machine Learning, Prediction, Adaboost. Confusion Matrix

(Received 2024-02-03, Accepted 2024-03-01, Available Online by 2024-03-08)

1. Introduction

Diabetes is a disease that threatens global public health but is challenging to detect early because of the lack of apparent symptoms. According to the World Health Organization (WHO), diabetes is ranked 9th as the deadliest disease in the world. In Indonesia itself, diabetes is a severe problem, ranking 7th as the country with the highest number of people with diabetes [1]. The impact of diabetes is comprehensive, starting from damaging vital organs such as the kidneys, eyes, and nerves to increasing the risk of heart disease and even death in mothers during childbirth. Delayed diagnosis often causes complications that lead to death before the patient realizes that they have diabetes [2].

The importance of early detection in preventing the destructive effects of diabetes drives the need for an effective prediction system [3]. By analyzing the supporting attributes of diabetes, prediction systems can be a more affordable and efficient option compared to visits to specialist doctors and laboratory tests. For this reason, machine learning methods, such as the Decision Tree algorithm, are often used in efforts to predict diabetes. In the field of prediction, several algorithms are often used in research related to machine learning, namely Decision Tree [4], Neural Network [5], Support Vector Machine, and Naive Bayes [6]. One of the machine learning algorithms for making predictions with the

highest level of accuracy, the most popular, which is easy for humans to understand and is often used, is the Decision Tree algorithm [7], [8].

Although the Decision Tree algorithm is effective, it has limitations in prediction accuracy, especially in the case of disease prediction [8]. Therefore, this research aims to increase prediction accuracy by integrating the Adaboost method into the Decision Tree algorithm. The Adaboost method is used to improve the performance of a single algorithm by forming several prediction models from training data [9].

This research aims to overcome the limited accuracy of diabetes prediction by integrating the Adaboost method into the Decision Tree algorithm. Through data analysis from the Kaggle dataset consisting of 520 records with 16 attributes, including age, gender, and other symptoms, this research evaluates the prediction performance using the Confusion Matrix. So, this research will provide an overview of the effectiveness of using the Decision Tree algorithm enhanced with the Adaboost method in predicting diabetes. The implications of the findings from this study will help improve understanding of the importance of predictive approaches in managing chronic diseases such as diabetes.

2. Methods

This research methodology outlines the steps that will be taken in the assessment process to achieve the stated research objectives. The Cross Industry Standard Process for Data Mining (CRISP-DM) data mining standardization model was chosen for this research methodology. CRISP-DM was selected because it is one of the most frequently used data mining methods. This study uses the Decision Tree classification model with the C4.5 algorithm, and the boosting is carried out using the Adaboost algorithm. Below is a picture of the flow of the stages of the research.

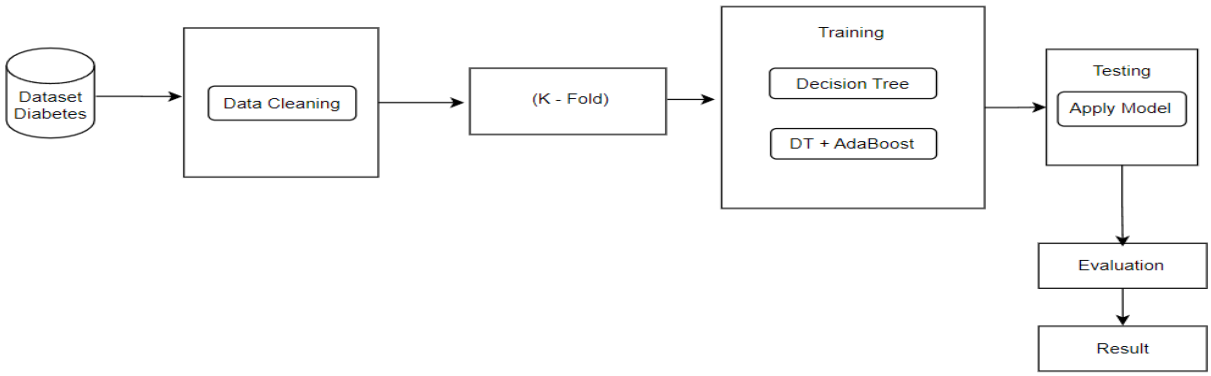


Figure 1. Research Methodology

2.1. Business Understanding

The first stage in CRISP-DM is understanding the business goals and needs from a business perspective. Based on the results of a literature study, it was found that diabetes is the deadliest disease, ranked 9th in the world. People who have diabetes have an increased risk of more severe and life-threatening health problems that can result in medical care costs, reduced quality of life, and increased mortality. With an estimated global prevalence of 9.3% in 2019, diabetes is a significant global public health problem, so computational analysis and disease prediction can help in diagnosis.

2.2. Data Understanding

This study uses public data from the Kaggle website with 520 data records. This dataset comprises 17 attributes, including age, gender, and symptoms that can influence a person's diabetes risk. This dataset's target field or label is diabetes disease status, with a negative value indicating not having the disease and a positive showing having the disease. The following dataset table is presented.

Table 1. Datasets

Age	Gender	Polydipsia	Sudden Weight Loss	…	Alopecia	Obesity	class
40	Male	Yes	No	…	Yes	Yes	Positive
58	Male	No	No	…	Yes	No	Positive
41	Male	No	No	…	Yes	No	Positive
45	Male	No	Yes	…	No	No	Positive
60	Male	Yes	Yes	…	Yes	Yes	Positive
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮

2.3. Data Preparation

After the data is collected, the next stage is the data preparation stage, including data cleaning. At the data cleaning stage, the initial data obtained is checked for missing/blank, noisy, and inconsistent data. If the data has missing values above 50%, the attributes can be ignored or deleted at the data cleaning stage. This can be done to overcome missing data values using the replace missing value method using the average contained in the return missing value. When processed, it produces a data pattern according to the dataset table, which explains that there is no empty/missing value or noisy or inconsistent data. With a total of 520 data.

2.4. Modelling

This stage directly involves Machine Learning to determine data mining techniques and algorithms. This research uses the Decision Tree classification model with the C4.5 algorithm, and the boosting is carried out using the Adaboost algorithm. This model will be implemented to see an immediate increase in accuracy.

1. Decision Tree C4.5

   Find tree roots. The selected attribute will be the basis for taking root, and the method for calculating the gain value for each attribute will be used; the highest gain value will be the first root. Before calculating the gain value for each attribute, calculate the entropy value first. The entropy value is calculated as follows:

$$Entropy(S) = \sum_{i=1}^{c} - p_i \log_2 p_i \quad (1)$$

   To calculate the Gain value, use the Equation:

$$Gain(S, A) = Entropy(S) - \sum_{i=1}^{n} \frac{|Si|}{|S|} * Entropy(Si) \quad (2)$$

2. Adaboost

   Then the Adaboost calculation is as follows:

   Initiation of weight values on training samples $$D^1(i) = \frac{1}{m}\ Untuk\ i = 1, \dots, m \quad (3)$$

   Calculate the training sample error $$€t = \sum Dt\,(i) \quad (4)$$

   Calculate the training sample weight values $$\alpha t = \frac{1}{2} \ln \left(\frac{1 - €t}{€t}\right) \quad (5)$$

After that, update the sample weight value for correct predictions

$$Dt + 1(i) = Dt(i)x\ \{^{-\alpha t} \quad (6)$$

After that, update the sample weight value for incorrect predictions

$$Dt + 1(i) = Dt(i)x\ \{^{\alpha t} \quad (7)$$

Output of final prediction

$$H(X) = sign\ (\sum_{t=1}^{T} \alpha t\ ht(X)) \quad (8)$$

2.5. Evaluation

Looking at for the performance level of the patterns created by the algorithm is a way to carry out this stage. The evaluation algorithm uses Matrix Confusion with arrangement for accuracy, precision, and recall values. It is possible to calculate this value:

$$\text{Accuracy} = \frac{\text{Number of correct classifications}}{\text{Total testing samples tested}}\ x\ 100\% \quad (9)$$

$$\text{Precision} = \frac{\text{True Positive}}{\text{True Positive+False Negative}}\ x\ 100\% \quad (10)$$

$$\text{Recall} = \frac{\text{True Positive}}{\text{True Positive+False Positive}}\ x\ 100\% \quad (11)$$

2.6. Deployment

Then, the evaluation stage is completed, where a specific and detailed assessment is made based on the results of a model. Therefore, implementation of all the models that have been created is carried out. Apart from what has been explained, adjustments are also made to the model to produce results following the initial objectives of the proposed CRISP-DM stage.

3. Results and Discussion

3.1. Modelling

At this stage, 520 records are used with data with 17 attributes used. The last attribute (Class) is the target class, so there are 16 data attributes. There are two treatments for modeling the dataset: data using the decision tree algorithm and dataset using the Decision Tree algorithm, which is optimized with Adaboost.

1. Decision Tree Modeling C4.5

   Modeling the C4.5 Decision Tree Algorithm begins by calculating the entropy value. After that, estimate the gain values of the ten attributes used to build the classification tree. The parent node is determined from the attribute with the highest gain value. From the parent node, a branch is created from the parent node category. Then, check whether there are any remaining attributes. If the condition is still there, repeat the process of calculating the entropy value. If the condition is not, continue building the C4.5 Decision Tree Algorithm, interpreting the results of the Decision Tree created, and calculating the accuracy value based on the confusion matrix. The results of calculating the entropy value using equation (1), the gain using the following equation (2):

Table 2. Calculation of Entropy and Gain Values

		Total (S)	Yes(Si)	No(Si)	Entropy	Gain
Age	23-45	203	100	103	0.9998	0.0385
	46-90	317	195	132	0.9612	

Gender	Male	328	148	180	0.9930	0.1452
	Female	192	170	22	0.5135	
Polyuria	Yes	258	230	28	0.4954	**0.3269**
	No	262	75	187	0.8638	
Polydipsia	Yes	233	220	13	0.3105	0.2802
	No	287	90	197	0.8972	
Sudden WL	Yes	217	187	30	0.5796	0.1452
	No	303	130	173	0.9853	
Weakness	Yes	305	225	80	0.8301	0.0614
	No	215	103	112	0.9986	
Polyphagia	Yes	237	191	46	0.7098	0.0948
	No	283	133	150	0.9974	
Genital	Yes	116	80	36	0.8935	0.0049
	No	404	240	164	0.9742	
Visual Blur	Yes	233	176	57	0.8026	0.0497
	No	287	145	142	0.9998	
Itching	Yes	253	153	100	0.9680	0.0004
	No	267	167	100	0.9540	
Irratability	Yes	124	110	14	0.5085	0.0864
	No	394	214	180	0.9945	
Delay heal	Yes	239	152	87	0.9460	0.0019
	No	281	165	116	0.9779	
Partial pares	Yes	224	180	44	0.7146	0.0917
	No	296	128	168	0.9867	
Muscle stif	Yes	195	130	65	0.9183	0.0303
	No	325	180	145	0.9916	
Alopecia	Yes	341	243	98	0.8653	0.0530
	No	179	79	100	0.99	
Obesity	Yes	88	53	35	0.9695	0.0181
Total(Class)		520	320	200	0.9612	

Based on the table above, it can be seen that of the 16 attributes used in this research, the Polyuria attribute has the highest gain value, namely 0.3269. This means that the polyuria attribute has the most significant influence in predicting diabetes. Then, the Polyuria attribute will be the root node.

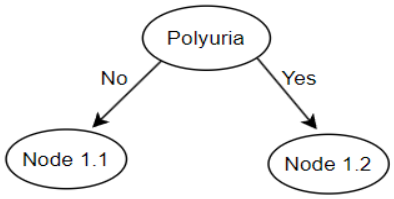


Figure 2. Node Root

Based on the root node above, the following nodes can be continued. Eliminate the previously selected attributes and repeat the calculation as at the beginning of the Entropy value, Information Gain, by choosing the largest Information Gain and making it the internal node of the tree. Repeat the calculation until all tree attributes have a class. Finally, a decision tree is produced as follows.

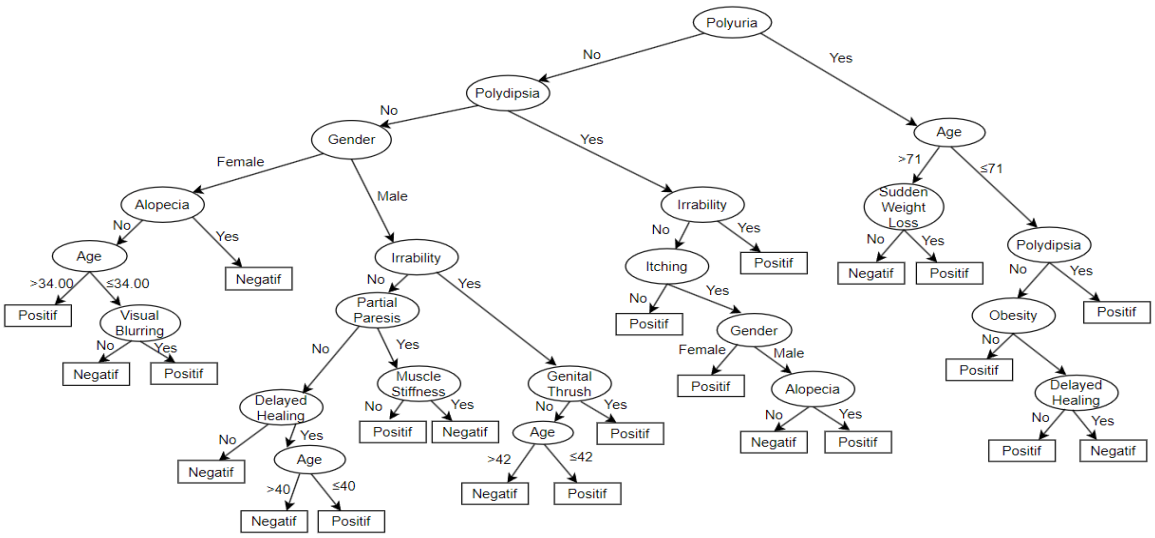


Figure 3. Decision Tree

Based on the decision tree image, a rule is formed as follows.
R1: **IF** Polyuria=No ^ Polydipsia=No ^ Gender=Female ^ Alopecia=Yes **THEN** Diabetes=Negatif
R2: **IF** Polyuria=No ^ Polydipsia=No ^ Gender=Female ^ Alopecia=No ^ Age= >34 **THEN** Diabetes=Positif
R3: **IF** Polyuria=No ^ Polydipsia=No ^ Gender=Female ^ Alopecia=No ^ Age= ≤34 ^ Visual Blurring=No **THEN** Diabetes=Negatif
R4: **IF** Polyuria=No ^ Polydipsia=No ^ Gender=Female ^ Alopecia=No ^ Age= ≤34 ^ Visual Blurring=Yes **THEN** Diabetes=Positif
R5: **IF** Polyuria=No ^ Polydipsia=No ^ Gender=Male ^ Irrability=No ^ Partial Paresis=No ^ Delayed Healing=No **THEN** Diabetes=Negatif

After the C4.5 Decision Tree Algorithm is built, the model is evaluated using a confusion matrix. The initial data is predicted based on the C4.5 Decision Tree Algorithm that has been created. It was found that 306 patients were declared to have diabetes, 184 patients were displayed not to have diabetes or were non-diabetic, and 16 patients with diabetes who were identified as non-diabetics were included in the "Type I Error." In contrast, non-diabetic patients who were identified as having diabetes, as many as 14 are included in the "Type II Error."

Table 3. Confusion Matrix Decision Tree

	true Positive	true Negative
pred. Positive	306	16
pred. Negative	14	184

The evaluation results of the C4.5 Algorithm model using the confusion matrix shown in Table 3 show an accuracy value of 94.23%. After that, modeling was carried out using the Adaboost method.

2. Adaboost Modeling
The initialization weight value of the data in the first iteration using Equation (3) with a maximum iteration of 10 is 0.00192. Data found that did not match the original class in the first iteration were 30 data. The next step is to calculate the research data error using Equation (4). The data error value

in the initial iteration is 0.0576. After calculating research data errors, the next step is to calculate the data weights using the Equation (5). The data weight value obtained was 1.3994. Using Equation (6), the data weights are updated in the first iteration where positive status is a positive result, while negative is not having diabetes. The weight of the data whose initial status was positive and was correctly predicted as positive and whose initial status was negative and correctly predicted as unfavorable was 0.000473. The weight of data whose initial status is positive and is expected to be hostile or vice versa is 0.007774.

Table 4. Hasil Update Bobot

No	Age	Gender	Polyuria	Polydipsia	Class	Class Prediction	Update Bobot
1	54.0	Female	Yes	Yes	Positive	Positive	0.000473
2	48.0	Female	Yes	Yes	Positive	Positive	0.000473
3	60.0	Male	Yes	Yes	Positive	Positive	0.000473
4	53.0	Male	Yes	Yes	Positive	Positive	0.000473
5	41.0	Male	Yes	Yes	Positive	Positive	0.000473
6	63.0	Male	Yes	Yes	Positive	Positive	0.000473
7	48.0	Female	No	No	Positive	Positive	0.000473
8	60.0	Female	Yes	Yes	Positive	Positive	0.000473
9	50.0	Female	No	Yes	Positive	Negative	0.007774
10	25.0	Female	No	No	Positive	Negative	0.007774
11	39.0	Female	Yes	Yes	Positive	Positive	0.000473
…	…	…	…	…	…	…	…

The calculation is repeated until the error value is at least 0.5 and the maximum iteration is reached. After that, the process can be stopped. Next, evaluate the model using the Adaboost method using a confusion matrix. Prediction data from the C4.5 Algorithm is predicted again based on the boosting results using the Adaboost method. It was found that 311 patients were declared to have diabetes, 195 patients were displayed not to have diabetes or were non-diabetic, and five patients with diabetes who were identified as non-diabetics were included in the "Type I Error." In contrast, non-diabetic patients who were identified as having diabetes as many as nine are included in the "Type II Error."

Table 5. Confusion Matrix Adaboost

	true Positive	true Negative
pred. Positive	311	5
pred. Negative	9	195

Based on the outcome of the evaluation of the C4.5, The algorithm model, after being boosted with a confusion matrix, uses the Adaboost method in Table 5; the accuracy value was 97.31%. This means there is an increase in the accuracy of the results obtained with the Decision Tree C4.5 algorithm. The accuracy value increased by 3% after boosting using the Adaboost method.
This is because the Adaboost method can handle samples that are difficult to predict by the decision tree model. Adaboost will improve its predictions on these samples by reducing the number of errors and increasing accuracy. The Decision Tree model that makes the best contribution to errors is given greater weight in Adaboost. Giving weight to the best model helps increase the strength of the ensemble and dominates the decisions of better decision tree models [10]. The Adaboost models produced at each iteration are combined into an ensemble model. The final prediction is made by taking the majority decision from all models. This can help overcome the weaknesses of the Decision Tree model.

4. Conclusion

This research emphasizes the importance of predicting diabetes early as an initial step in preventing the negative impacts that may arise due to this disease. With global prevalence continuing to increase, early detection is crucial to reduce the risk of potentially fatal complications. The research results show that using the Adaboost method to improve the accuracy of diabetes predictions is very effective. The integration of Adaboost with the Decision Tree algorithm increased prediction accuracy from 94.23% to 97.31%. This shows that ensemble learning can be an effective solution to improve the performance of a single algorithm. The findings of this study effectively highlight the potential of the Adaboost method in improving the accuracy of diabetes prediction. By improving predictions on samples that are difficult for Decision Tree models to predict, Adaboost manages to reduce the number of errors and increase overall accuracy. This suggests that Adaboost could be helpful in early disease detection efforts. With increased prediction accuracy, healthcare can provide patients with more precise diagnoses and earlier treatment. This can help in reducing the risk of potentially fatal complications due to diabetes. This research still does not have any treatment for exploration that can further explore the influence of Adaboost parameters, such as the number of iterations (T) that can influence the results of Adaboost [11].

References

[1] Kementerian Kesehatan RI., "Infodatin tetap produktif, cegah, dan atasi Diabetes Melitus 2020," *Pusat Data dan Informasi Kementerian Kesehatan RI*. pp. 1–10, 2020.

[2] S. Ucha Putri, E. Irawan, F. Rizky, S. Tunas Bangsa, P. A. -Indonesia Jln Sudirman Blok No, and S. Utara, "Implementasi Data Mining Untuk Prediksi Penyakit Diabetes Dengan Algoritma C4.5," *Januari*, vol. 2, no. 1, pp. 39–46, 2021.

[3] M. R. Firmansyah and Y. P. Astuti, "Stroke Classification Comparison with KNN through Standardization and Normalization Techniques," *ASSET*, vol. 6, no. 1, pp. 1–8, 2024.

[4] P. Songthung and K. Sripanidkulchai, "Improving type 2 diabetes mellitus risk prediction using classification," *2016 13th Int. Jt. Conf. Comput. Sci. Softw. Eng. JCSSE 2016*, 2016, doi: 10.1109/JCSSE.2016.7748866.

[5] M. Komi and X. Zhang, "Application of Data Mining Methods in Diabetes Prediction Messan," no. S Ix, pp. 1006–1010, 2017.

[6] R. S. Raj, D. S. Sanjay, M. Kusuma, and S. Sampath, "Comparison of Support Vector Machine and Naïve Bayes Classifiers for Predicting Diabetes," *1st Int. Conf. Adv. Technol. Intell. Control. Environ. Comput. Commun. Eng. ICATIECE 2019*, pp. 41–45, 2019, doi: 10.1109/ICATIECE45860.2019.9063792.

[7] L. Amini *et al.*, "Prediction and control of stroke by data mining," *Int. J. Prev. Med.*, vol. 4, pp. S245–S249, 2013.

[8] T. Kansadub, S. Thammaboosadee, S. Kiattisin, and C. Jalayondeja, "Stroke risk prediction model based on demographic data," *BMEiCON 2015 - 8th Biomed. Eng. Int. Conf.*, pp. 3–5, 2016, doi: 10.1109/BMEiCON.2015.7399556.

[9] T. Asra, A. Setiadi, M. Safudin, E. W. Lestari, N. Hardi, and D. P. Alamsyah, "Implementation of AdaBoost Algorithm in Prediction of Chronic Kidney Disease," *2021 7th Int. Conf. Eng. Appl. Sci. Technol. ICEAST 2021 - Proc.*, pp. 264–268, 2021, doi: 10.1109/ICEAST52143.2021.9426291.

[10] J. Ha, M. Kambe, and J. Pe, *Data Mining: Concepts and Techniques*. 2011. doi: 10.1016/C2009-0-61819-5.

[11] A. H. Yunial, "Optimization Analysis of Support Vector Machine Classification Algorithms, Decision Trees, and Neural Networks Using Adaboost and Bagging," *J. Inform. Univ. Pamulang*, vol. 5, no. 3, p. 247, 2020.

Advance Sustainable Science, Engineering and Technology (ASSET)
Vol. 6, No.2, April 2024, pp. 0240208-01 ~ 0240208-08
ISSN: 2715-4211 DOI: https://doi.org/10.26877/asset.v6i2.18163

Identification of Landslide Prone Areas with Schlumberger Configuration Geoelectric Method, Kalongan Village, East Ungaran in 2023

Zeni Rahmawati[1*], Thaqibul Fikri Niyartama[1], Nugroho Budi Wibowo[1,2], Andi[1]

[1]Faculty of Science and Technology, Universitas Islam Negeri Sunan Kalijaga, Jl. Laksda Adisucipto, Papringan, Caturtunggal, Kec. Depok, Kabupaten Sleman, Daerah Istimewa Yogyakarta 55281

[2]BMKG, Stasiun Geofisika Sleman, D.I Yogyakarta

*zenirahmawati28@gmail.com

Abstract. Landslides are one of the natural disasters that occur in Semarang Regency. The landslide event in Kalongan Village, East Ungaran Subdistrict took place in 2022 to 2023. The dimension of the landslide is large and has the potential to cause damage to infrastructure and settlements, mitigation measures are needed. One of the mitigation measures is to identify the lithology of the landslide area and its surroundings. This study aims to determine the lithology of the landslide area and its surroundings using the Schlumberger configuration geoelectric method. Geoelectric acquisition using a set of naniura geoelectric, with a line variation of 120 to 300 meters. The number of measurement lines is 6 VES located in the landslide zone in the Kaligetas Formation. The results showed that the avalanche zone is composed of Tufan Sandstone with a resistivity of 0.17 Ωm to 334.9 Ωm, Tuff with a resistivity of 10.03 to 26.96 Ωm, and Lava Flow with a resistivity of 1161.91 Ωm to 3040.05 Ωm. Avalanche materials were identified to be associated with lithologies such as tufan sandstone and tuff with a thickness between 5 to 50 m.

Keywords: Lithology, Landslide, Geoelectric, Schlumberger Configuration, Ungaran.

(Received 2024-01-15, Accepted 2024-02-11, Available Online by 2024-03-08)

1. Introduction

Natural disasters are defined as disasters caused by natural events, such as tsunamis, volcanic eruptions, earthquakes, droughts, floods, hurricanes and landslides. One example of a natural disaster is a landslide. Landslide has a definition of moving rock, detritus, or soil caused by the force of gravity [1]. Landslides also have another definition which is the movement of soil or rock down a slope due to local geological factors, groundwater conditions, extreme weather, earthquakes, or other factors [2]. Factors affecting landslides consist of internal factors (slope building, rock lithology, geological structure, and sliding plane) and external factors (high rainfall and earthquakes). One of the factors that can determine the

occurrence of landslides is the lithology of the area. Lithology is included in the characteristics or properties of rocks that consist of mineral composition, structure, and forming materials.

Based on data from the Regional Disaster Management Agency (BPBD) of Semarang Regency in 2022, the area of Semarang Regency is categorized as at risk of landslides with a total incidence of 108.The area includes Bringin Subdistrict, Banyubiru Subdistrict, Getasan Subdistrict, Sumowono Subdistrict, Suruh Subdistrict, Pringapus Subdistrict, West Ungaran Subdistrict, and East Ungaran Subdistrict [3]. Landslides that occurred in East Ungaran subtodistrict resulted in the disconnection of the connecting road between East Ungaran and Mranggen. Landslide movement continues to this day, oriented from the south to the north. Landslide movement has entered residential areas, this can be seen in **Figure 1**.

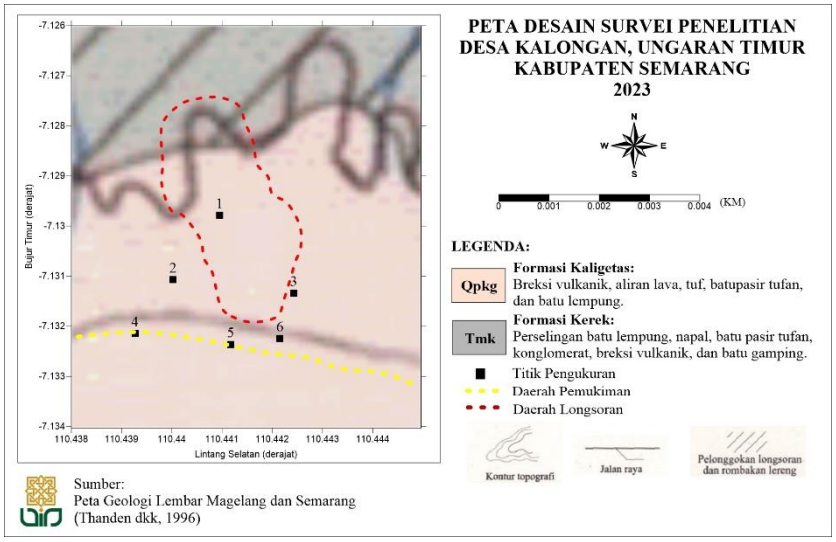
(**Figure 1.** Landslide Area)

Landslides can be minimized by mitigation. Mitigation has a definition of the stages of how to reduce disaster risk in terms of physical development or the ability to deal with disasters. One of the mitigations carried out to minimize damage and casualties is by using geoelectric methods. There are various types of geoelectric method configurations, one of which is the Schlumberger configuration. The Schlumberger configuration geoelectric method can be used to identify rock layers and layer thickness. This can be done because rocks have electrical properties that depend on the constituent minerals [4]. The Schlumberger configuration has the advantage of being able to detect the presence of inhomogeneous rock layers.

Research conducted in this area has been conducted by Lestari (2019) and Zulaikhah (2020) covering Semarang Regency using the Geographic Information System (GIS) method, and Vijaya (2022) covering Semarang Regency using the Fellinius method. There is no research in this area that uses geoelectric methods, lithological analysis, and research in Kalongan Village specifically. Thus, this study aims to identify lithology using geoelectric methods in Kalongan Village, East Ungaran. This research is important because the landslide area is densely populated and continues to experience movement.

2. Methods

The research was conducted in Kalongan Village, East Ungaran with the Schlumberger configuration geoelectric method. The Schlumberger configuration geoelectric method has advantages and disadvantages, one example of its advantage is that it can detect the non-homogeneity of rock layers on the surface by comparing resistivity values and its weakness is that it requires current sending equipment that has a very high direct current (DC) voltage to overcome small potential electrode voltage readings. This study used 6 VES located in the landslide area with varying line lengths, short lines of 120 m and 300 m. The varying track lengths are due to the different areas of the study where there is soil on the surface along the track. The data were collected using a set of naniura geoelectric tools. Data collection uses a set of naniura geoelectric tools. Data processing used Microsoft Excel software, Progress to display 2D modeling, and Rockwork 16 to display 3D modeling. The research survey design map can be seen in **Figure 2**.



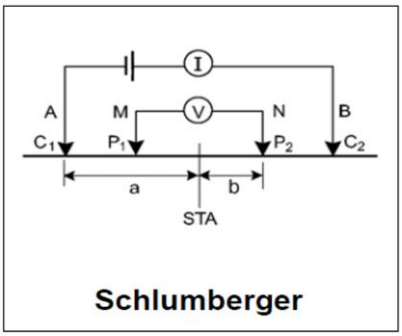
(**Figure 2.** Survey Design Map of Data Collection of Schlumberger Configuration Geoelectric Method with 6 VES)

2.1. Regional Geology

The study area is located in the kaligetas formation. This formation is composed of volcanic breccia, lava flows, tufan sandstone and claystone. Breccia flows and lava are interbedded with fine to coarse lava and tuff. The lower part is claystone that contains mollusks and tufan sandstone. Weathered volcanic rocks have a brown to reddish color by forming large chunks and a thickness of 50to200 m.

2.2. Geoelectric Method

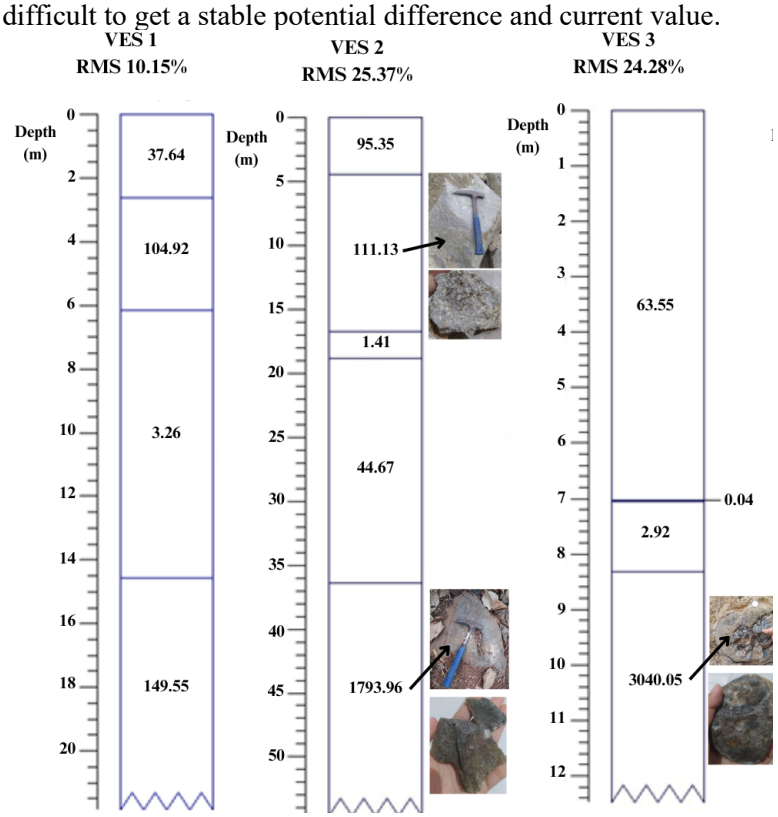
Geoelectric method is a geophysical method that studies the nature of electric currents in the earth and detects electric currents on the earth's surface. This method includes measurements of naturally occurring potentials and measurements of currents generated by the injection of electricity into the earth. There are various types of geoelectric methods, one of which is the resistivity geoelectric method. The type resistance method can describe the rocks below the ground surface. This can be seen by the difference in resistivity values in each rock. The working principle of the geoelectric method is by injecting electric current (I) into the subsurface of the earth using electrodes, so that it will produce a potential difference value (ΔV). Based on the survey objectives, the resistivity geoelectric method is divided into two, one of which is the sounding resistivity geoelectric method (Schlumberger configuration). The sounding resistivity geoelectric method is used to determine the variation of resistivity values vertically. The Schlumberger configuration uses four electrodes, two current electrodes and two potential electrodes. This can be shown in **Figure 3**.



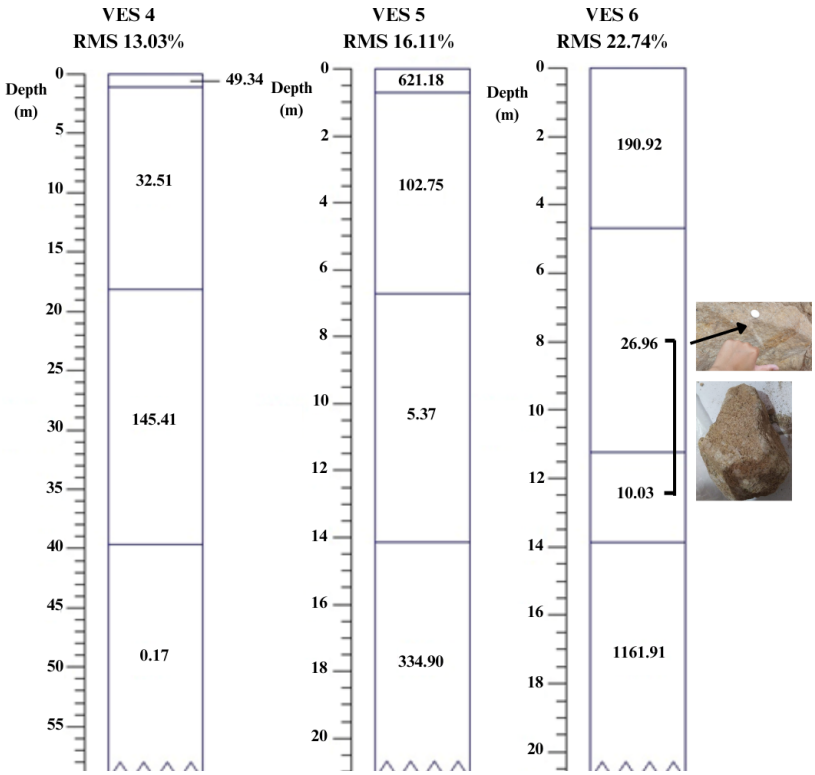
(**Figure 3.** Electrode arrangement of Schlumberger Configuration Geoelectric Method**)**

3. Results and Discussion

Measurements were taken in the landslide deposit area, the west, east and south sides of the landslide, as well as houses close to the landslide. The measurement location aims to determine whether the research area will continue to occur landslides by knowing the lithological constituents in the area. The results of data processing using progress software obtained resistivity value results that can be seen in **Figure 4** and **Figure 5**. VES 1 has a line length of 140 m with the measurement location right in the landslide sediment area. The results of the VES 1 data processing, getting a depth of 21.5 m with a resistivity value of 3.26 to 149.55 Ωm has an error of 10.15%. VES 2 has a line length of 300 m with the measurement location on the west side of the landslide. The results of the VES 2 data processing, getting a depth of 54 m with a resistivity value range of 1.41 to 1793.98 Ωm has an error of 25.37%. VES 3 line has a length of 240 m with the measurement location on the East side of the landslide. The results of the VES 3 data processing, getting a depth of 12.6 m with a resistivity value range of 0.04 to 3040.05 Ωm has an error of 24.28%. VES 4 has a line length of 300 m with the measurement location on the south side of the landslide close to residents' houses and road access. The results of the VES 4 data processing, getting a depth of 59 m with a resistivity value range of 0.17 to 145.41 Ωm has an error of 13.03%. VES 5 has a line length of 120 m with the measurement location on the south side close to residential houses. The results of the VES 5 data processing, getting a depth of 21 m with a resistivity value range of 5.37 to 621.18 Ωm has an error of 16.11%. VES 6 has a line length of 300 m with the measurement location on the south side of the landslide. The results of the VES 6 data processing, getting a depth of 20.5 m with a resistivity value range of 10.03 to 1161.91 Ωm has an error of 22.74%. The results of data processing that produce the smallest error are in VES 1 and the largest error in VES 2. This is because in VES 1 it is easier to stick the electrode because it is in the avalanche sediment area compared to VES 2 because it is in an area close to people's homes and access roads for pedestrians and motorbikes, so that the electrode that is stuck to the ground is difficult even though it has moved the electrode and it is difficult to get a stable potential difference and current value.



(**Figure 4.** Data Processing Results of VES 1 to 3 Resistivity Log Value and Depth)



(**Figure 5.** Data Processing Results of VES 4to 6 Resistivity Log Value and Depth)

The resulting resistivity value will be interpreted into several types of rocks. The resistivity value for rock interpretation refers to the 1990 Telford reference and field conditions. The resistivity value in the 1990 Telford reference can be seen in **Table 1** and the resistivity value of the research results can be seen briefly in **Table 2**.

Table 1. Resistivity values Telford, 1990.

No.	Litologi	Resistivitas (Ωm)
1	Batu Pasir	1 - 6.4 x 10^8
2	Tuff	2 x 10^3 - 10^5
3	Aliran Lava	10^2 - 5 x 10^4

The interpretation of the data processing that has been done, obtained in VES 1 is represented by soil (49.34 Ωm) and tuff sandstone (104.92, 3.26 and 149.55 Ωm). The interpretation of VES 2 is soil (95.35 Ωm), tuff sandstone (111.13, 1.41 and 44.67 Ωm), and lava flow (1793.96 Ωm). The interpretation of VES 3 is soil (63.55, 0.04 and 2.92 Ωm) and lava flow (3040.05 Ωm). This line is thought to have rocks that absorb water, so that at a depth of 7 m has a low resistivity value. The interpretation of VES 4 is soil (49.34 and 32.51 Ωm) and tuff sandstone (145.41 and 0.17 Ωm). This line is suspected to have rocks that absorb water, so that at a depth of 45 to 59 m has a low resistivity value. The interpretation of VES 5 is soil (621.18, 102.75, 5.37 Ωm) and tuff sandstone (334.90 Ωm). The interpretation of VES 6 is soil (190.92 Ωm), tuff (26.96 and 10.03 Ωm), and lava flow (1161.91 Ωm). The results of data interpretation, obtained rocks that are almost found in all lines are tuff sandstone. While lava flows are only found in VES 2, VES 3, and VES 6. Meanwhile, tuff is only found in VES 6.

Table 2. Resistivity Value of Research Results

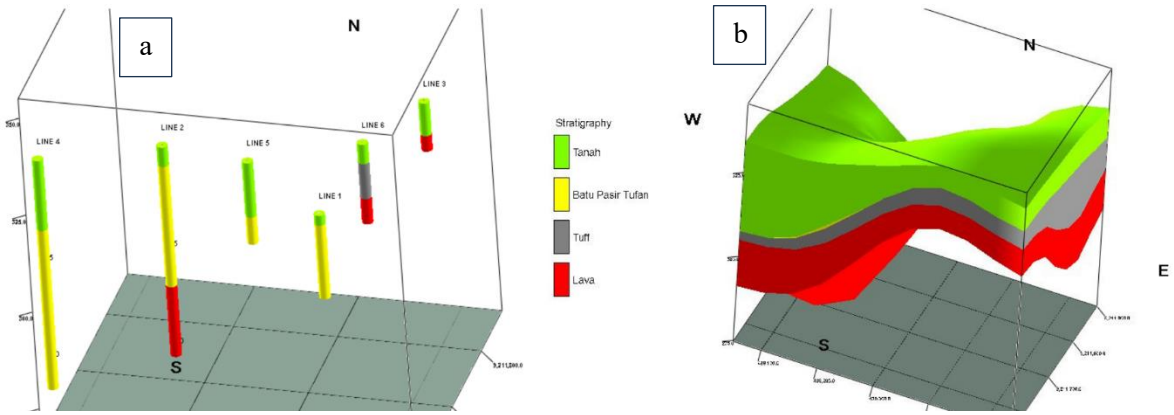
No.	Lithology	Resistivity (Ωm)
1	Top Soil	0.04 to 621.18
2	batu pasir tufan	0.17 to 334.9
3	tuff	10.03 to 26.96

4	lava	1161.91 to 3040.05

Lithologic conditions are one of the causes of landslides. One of the factors that can cause landslides is local groundwater, in the research area there is a water flow, this can be seen in **Figure 6**. Rocks that will experience landslides in this research area are tuff sandstone, claystone, and tuff. This is because tuff sandstone and tuff have low bearing capacity according to their grain size, if these rocks absorb water beyond their capacity, landslides will occur because the rocks are unable to withstand the water. Clay stones have impermeable properties and will expand when exposed to water, so clay stones are prone to landslides. Tuff sandstone, tuff, and claystone are included in sedimentary rocks. Sedimentary rocks generally have a lower resistivity value compared to igneous rocks, this is because each rock has different shape characteristics to absorb water.

(**Figure 6.** Puddles and Water Flow in the Avalanche Area)

The results of the data interpretation that has been done, can be produced 3D modeling using rockwork 16 software. The results of this modeling show the depth of each research line along with the depth of different rock types, so that it can be known visually the depth of the rock that has been interpreted. 3D modeling was made to show that each measured traverse consists of different rock layers. this can be seen in **Figure 7**. Multilog 3D modeling resulted in depths of less than 25 m at VES 1 and VES 5 not yet represented igneous rocks, namely lava flows. This is because the length of the measurement line is too short in the area. While on VES 3 and 6 with a depth of less than 25 m represented igneous rocks, namely lava flows, the short line length does not affect this area. The depth of VES 4 reaches 50 m, although this area has a deeper depth than the other lines. The measurement points in this area have not yet been represented by igneous rocks, thus requiring a longer measurement trajectory. The 3D stratigraphic modeling shows that the landslide that occurred in the study area moved from the south to the north. This can be seen in the first line measurement point which has a lower elevation than the other line measurement points. The East side of the avalanche has a higher elevation than the West and South sides of the avalanche, in this model the soil (top soil) is more than other rocks. Multilog 3D modeling can be a reference that the research area will be more prone to landslides seen from the interpreted rocks and stratigraphic 3D modeling that clearly shows the basin in the landslide area and the depth between rock layers.



(**Figure 7.** a.3D Multilog modeling, b.3D Stratigraphic modeling)

4. Conclusion

Based on the research that has been done, it can be concluded that the resistivity value in each track is different so that it can be interpreted into several types of rocks. VES 1, VES 4, and VES 5 are composed of the same rocks, namely soil (top soil) and tuffaceous sandstone. VES 2, VES 3, and VES 6 have the same rock composition, namely soil and lava flow, which distinguishes tuff in VES 6 and tuff sandstone in VES 2. The research conducted resulted in the research area being represented by dominant sedimentary rock, namely tuff sandstone. These rocks will be more easily weathered than igneous rocks, namely lava flows. Even this research area has a fairly heavy water flow in the rock layer on the South and East sides. This makes the area prone to landslides. The suggestion for research in Kalongan Village, East Ungaran is to use other methods to analyze the rock layers in the area to be more accurate.

References

[1] F. V. De Blasio, *Introduction to the physics of landslides: lecture notes on the dynamics of mass wasting*. Dordrecht ; New York: Springer, 2011.

[2] S. Szalai, K. Szokoli, M. Metwaly, Z. Gribovszki, dan E. Prácser, "Prediction of the location of future rupture surfaces of a slowly moving loess landslide by electrical resistivity tomography," *Geophysical Prospecting*, vol. 65, no. 2, hlm. 596–616, Mar 2017, doi: 10.1111/1365-2478.12421.

[3] S. Lestari, A. L. Nugraha, dan H. S. Firdaus, "Jurnal Geodesi Undip," vol. 8, no. 1, 2019.

[4] M. Farhati dan M. S. Rosid, "Identifikasi Bidang Gelincir dengan Metode Geolistrik Tahanan Jenis 2 Dimensi di Daerah Keranggan, Tangerang Selatan," *Positron*, vol. 12, no. 1, hlm. 1, Mei 2022, doi: 10.26418/positron.v12i1.53584.

[5] "telford-geldart-sheriff-applied-geophysics.pdf."

[6] F. Faizana, A. L. Nugraha, dan B. D. Yuwono, "Jurnal Geodesi Undip," vol. 4, 2015.

[7] R. M. Vijaya dan B. A. Arizaky, "ANALISA STABILITAS LERENG LONGSOR DI UNGARAN TIMUR KABUPATEN SEMARANG".

[8] V. Arsanti, R. S. Kharisma, dan S. Arfianto, "Analysis Spatial Pattern Garbage Bank using Web Geographic Information System in Yogyakarta City," *ASSET*, vol. 5, no. 1, hlm. 0230102, Apr 2023, doi: 10.26877/asset.v5i1.15139.

[9] K. Tjandra, *Empat bencana geologi yang paling mematikan*, Cetakan pertama. Yogyakarta: Gadjah Mada University Press, 2017.

[10] "[Gary_Mavko,_Tapan_Mukerji,_Jack_Dvorkin]_The_Rock(Bookos.org).pdf."

[11] B. Usman, R. H. Manrulu, A. Nurfalaq, dan E. Rohayu, "IDENTIFIKASI AKUIFER AIR TANAH KOTA PALOPO MENGGUNAKAN METODE GEOLISTRIK TAHANAN JENIS KONFIGURASI SCHLUMBERGER," *FLUX*, vol. 14, no. 2, hlm. 65, Okt 2017, doi: 10.20527/flux.v14i2.4091.

[12] N. Muhlisah, R. A. Magfirah, I. N. Mutiah, dan A. Pratama, “Analisis Litologi Bawah Permukaan Teluk Liukang Menggunakan Metode Seismik Refraksi,” vol. 01, no. 01, 2020.
[13] E. S. Putriutami dan U. Harmoko, “INTERPRETASI LAPISAN BAWAH PERMUKAAN DI AREA PANAS BUMI GUNUNG TELOMOYO, KABUPATEN SEMARANG MENGGUNAKAN METODE GEOLISTRIK RESISTIVITY KONFIGURASI SCHLUMBERGER,” vol. 3, no. 2, 2014.
[14] D. D, B. Nurlaksito, dan B. Legowo, “Identifikasi Bidang Gelincir Pemicu Bencana Tanah Longsor Dengan Metode Resistivitas 2 Dimensi Di Desa Pablengan Kecamatan Matesih Kabupaten Karanganyar,” *Indonesian J Appl Phys*, vol. 2, no. 02, hlm. 51, Mei 2016, doi: 10.13057/ijap.v2i02.1292.
[15] A. Masyhuri, H. Purnaweni, A. R. Herawati, dan B. P. Priyadi, “Kolaborasi Antar Stakeholders Dalam Manajemen Bencana Tanah Longsor Di Kota Semarang,” *j.Educ.hum.soc.sciences*, vol. 4, no. 2, hlm. 854–862, Okt 2021, doi: 10.34007/jehss.v4i2.759.
[16] M. A. Syam, K. Sasmito, dan N. N. Adlina, “GEOLOGI DAN PENGARUH LITOLOGI TERHADAP BENTUK MORFOLOGI DAERAH BANGUN REJO KECAMATAN TENGGARONG SEBERANG KABUPATEN KUTAI KARTANEGARA,” vol. 1, no. 1, 2018.
[17] Y. Arman, “Identifikasi Struktur Bawah Tanah di Kelurahan Pangmilang Kecamatan Singkawang Selatan Menggunakan Metode Geolistrik Resistivitas dan Inversi Lavenberg - Marquardt,” *positron*, vol. 2, no. 1, Mei 2012, doi: 10.26418/positron.v2i1.2001.
[18] M. Kadri dan T. F. Sudarma, “Penentuan Struktur Bawah Permukaan Daerah Geothermal Menggunakan Metode Geolistrik di Desa Penen Kecamatan Biru Biru Kabupaten Deli Serdang,” *TGEO*, vol. 8, no. 1, hlm. 1, Nov 2019, doi: 10.24114/tgeo.v8i1.13914.
[19] D. Trisnawati, Najib, dan A. S. Hidayatillah, “The Relationship of Lithology with Landslide Occurrences in Banyumanik and Tembalang Districts, Semarang City,” *IOP Conf. Ser.: Earth Environ. Sci.*, vol. 1047, no. 1, hlm. 012026, Jul 2022, doi: 10.1088/1755-1315/1047/1/012026.
[20] M. I. R. Marani, N. Najib, dan R. K. Ali, “Penentuan Zona Gerakan Tanah dan Analisis Kemantapan Lereng di Kecamatan Klego, Kabupaten Boyolali, Jawa Tengah,” *Jurnal Geosains dan Teknologi*, vol. 1, no. 3, hlm. 89, Nov 2018, doi: 10.14710/jgt.1.3.2018.89-98.

Advance Sustainable Science, Engineering and Technology (ASSET)
Vol. 6, No.2, April 2024, pp. 0240209-01 ~ 0240209-09
ISSN: 2715-4211 DOI: https://doi.org/10.26877/asset.v6i2.18307

Enhancing Web Server Security against Layered Cyber Threats in Healthcare

Muhammad Fajar Rifai, Syaiful Hendra*, Hajra Rasmita Ngemba, Ryfial Azhar, Rahmah Laila

Information Technology Department, Faculty of Engineering, Tadulako University, Jl.Soekarno Hatta No KM 9 Palu 94148, Central Sulawesi, Indonesia

*syaiful.hendra.garuda@gmail.com

Abstract. *Information technology plays an important role in improving operational efficiency at Torabelo Hospital. The server system in use today faces security and optimization challenges. This research analyzes the impact and recommends solutions to improve server security and optimization. The findings show that the server system is vulnerable to various types of attacks and performance degradation. This can negatively impact hospital operations and put patients at risk. The recommended solution is to implement Squid as reverse proxy, WAF (Web Application Firewall), and Snort as IDS (Intrusion Detection System). System testing showed that this solution successfully detected and prevented various common attacks. This research provides insights to health IT professionals to improve the security and performance of their server systems and improve healthcare services to patients at Torabelo Hospital.*

Keywords: Hospital, Security, Squid, Web Application Firewall, Healthcare

(Received 2024-01-31, Accepted 2024-02-13, Available Online by 2024-03-08)

1. Introduction

In today's digital age, information technology is at the heart of operational efficiency in many sectors, including the healthcare industry. As an integral part of its healthcare system, Torabelo Hospital has adopted a server system to manage patient data, organize medical procedures, and provide quick access to critical information. However, these advancements pose serious server security and optimization challenges that impact the hospital's operations. Security issues that could result in sensitive patient data leakage or delays in clinical information delivery due to lack of server optimization require in-depth understanding and appropriate solutions[1]. Torabelo Hospital's server security issues have drawn attention as potentially serious. Security threats such as hacking, malware, and even unauthorized access will result in losses and compromise the privacy and security of patients' personal information [2,3]. Server security in healthcare is a key issue when protecting sensitive patient data. Various cyber-attacks such as malware, phishing, and DDoS can compromise servers and cause data leakage. Unauthorized access through weak passwords, excessive privileges, and stolen credentials can compromise patient privacy [16, 17]. Patient data security is an important aspect of modern healthcare systems. Protecting

sensitive medical records requires ensuring various aspects such as network security, access, data encryption, and user identity management. Such protection is important to maintain patient privacy, improve healthcare quality, and comply with applicable regulations[18]. Effective information security risk management is essential to prevent or minimize the impact of unwanted incidents. This can be achieved by identifying, analyzing and evaluating risks and implementing appropriate risk mitigation measures [19].

Information security focuses on three main pillars: confidentiality, integrity, and availability. These pillars are reinforced by authentication, authorization, auditing, and non-denial. Proper implementation will build trust and improve efficiency in the digital age [20]. Information security such as protecting information and devices in the workplace. Minimizing damage from various threats. Its aspects include privacy, identification, authentication, authorization and accountability. Its presence ensures a smooth and secure workflow [21]. Information security is an important consideration when creating an integrated clinical environment. Development should be based on a fundamental system that ensures confidentiality of patient data, integrity of information, controlled access, and accountability of all actions. The implementation of these systems creates a strong fortress that protects sensitive information, ensures smooth workflow, and builds trust between patients and medical staff. Information security is not the responsibility of one party alone; all elements of the healthcare ecosystem must work together to create safe and reliable services [22]. Information security is like a solid fortress that protects a company's valuable assets from various threats. Its presence minimizes losses and helps the company achieve its goals with strong internal and external control systems. Its implementation demonstrates an organization's commitment to data security and stakeholder trust [21].

In addition, lack of server optimization slows down data access and processing, which are critical factors in a medical environment that can impact timely diagnosis and treatment of patients. Therefore, it is important to understand the impact of these security issues and lack of optimization on overall hospital operations and develop appropriate strategies to address them. The purpose of this study is to document and analyze in detail the security challenges faced by Torabelo Hospital's servers and to determine the impact of the lack of optimization of medical servers. By understanding the causes of these issues, this research provides valuable insights for healthcare IT professionals and other stakeholders to take effective steps in improving the security and optimization of their server systems, thus ultimately expected to support the provision of sound medical services. It is also expected to provide better and more efficient services for patients at Torabelo Hospital.

2. Methods

2.1. Stages of research

The hospital's IT infrastructure consists of a TCP/IP network with Internet access, servers with Ubuntu operating systems, EMR and HIS applications, and hardware such as servers, firewalls, routers, and workstations. The use of Ubuntu operating system on the hospital's server improves the security and stability of the IT system. The EMR application helps hospitals store and manage patient records electronically, thus improving the efficiency and accuracy of medical documentation. The HIS integrates various administrative and clinical functions of the hospital and improves the efficiency and effectiveness of hospital operations. High-speed internet connectivity ensures smooth operation of the IT system and access to patient data. At this stage of the research, the researcher will outline the steps taken in this study. The following steps are carried out as follows :

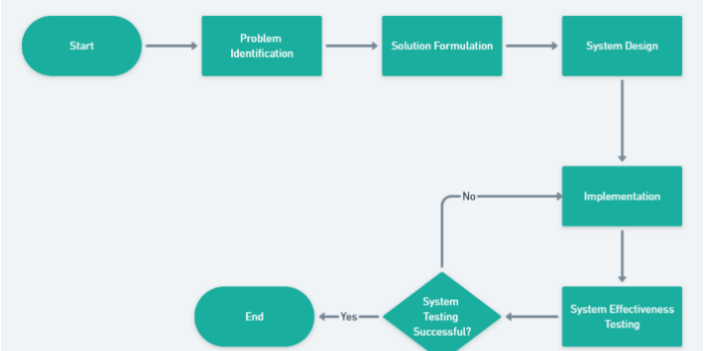


Figure 1. Research Flow Chart

During the problem identification phase, the entire web system requires additional protection against potentially harmful cyber attacks such as Deface, SQL injection, XSS, and DDoS. In addition, it requires a system that can monitor and detect suspicious activity[4,5]. The formulated solution involves using Squid reverse proxy as the first layer of defense[3,6]. Squid acts as an intermediary that directs incoming traffic to the Apache web server. Additionally, the solution includes the implementation of a web application firewall (WAF) to filter and prevent attacks such as tampering, SQL injection, and XSS on incoming traffic. Properly configured IDS (Intrusion Detection System) and IPS (Intrusion Prevention System) settings are also required for early detection and further prevention[7-9]. The system design considers the Apache configuration as the main web server and is configured to handle requests after passing through the Squid reverse proxy[10,11]. The Squid configuration filters incoming traffic and forwards it to Apache for processing. Squid's built-in WAF implementation to monitor and prevent the attacks mentioned above. IDS and IPS are installed at the network layer to monitor traffic, detect suspicious activity, and automatically prevent attacks if detected. The implementation includes configuring Apache as a web server, setting up Squid reverse proxy, installing and configuring WAF to prevent certain attacks, and adding IDS to your network environment to monitor and protect traffic flowing through your system, and installing IPS[12,13].

2.2. System Design

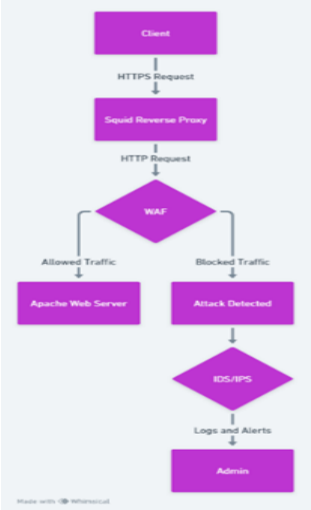


Figure 2. System Design

A system that combines Apache as a web server and Squid as a reverse proxy and designs a configuration that sets up Squid as the first layer before forwarding access to the Apache web server to secure the web infrastructure[2]. In this configuration, Squid acts as the initial gateway that receives requests from users. Squid completes some additional validation and processing phases before

forwarding the request to Apache. This allows the use of access control policies and caching, which improves security by optimizing requests to internal servers before they reach the main web server[14].

The application of Web Application Firewall (WAF) when designing this system is an important step to protect web applications from various malicious attacks such as Deface, SQL Injection, XSS, and DDoS attacks. WAF works by analyzing HTTP requests and filtering incoming content to the web application to ensure that the received data does not contain malicious payloads or programmatic attacks. With proper configuration[15], WAF provides a strong and adaptable layer of defense against evolving security threats by setting access policies and blocking attacks before they reach the web application[9].

The implementation of an Intrusion Detection System (IDS) and Intrusion Prevention System (IPS) into the system strengthens the defense layer against attacks. IDS monitors and analyzes network traffic to detect suspicious behavior patterns and attacks in progress. IPS, on the other hand, acts as an active defense against attacks by shutting down or blocking traffic identified as potential threats. With IDS and IPS, the system becomes more responsive in detecting, responding to, and protecting your infrastructure from various malicious attacks[5,10].

3. Results and Discussion.

3.1. Squid Reverse Proxy Configuration

The first step is to install Squid on the Ubuntu server, then configure the reverse proxy as the first layer before switching to the Apache web server in "Squid. Conf.

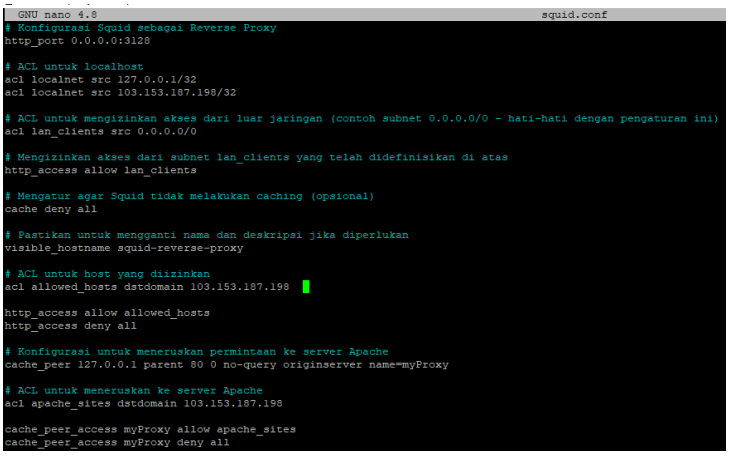


Figure 3. Squid Reverse Proxy Configuration

3.2. Web Application Firewall (WAF) Configuration

The steps taken in the WAF configuration are to install the ModSecurity "Library" and then activate "ModSecurity" in the Apache configuration at the end of the file.

LoadModule security2_module /usr/lib/apache2/modules/mod_security2.so

Then, define the security rules in ModSecurity.conf.

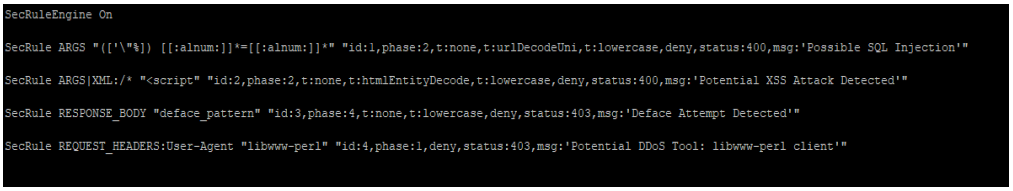


Figure 4. Rules on WAF

Add "/etc/modsecurity/webproxy_rules.conf" at the end of the "ModSecurity.conf" block and configure the VirtualHost block to enable ModSecurity.

```
<IfModule security2_module>
 SecRuleEngine On
 SecRule ARGS "([\'\"%]) [[:alnum:]]*=[[:alnum:]]*" "id:1,phase:2,t:none,t:urlDecodeUni,t:lowercase,deny,status:400,msg:'Possible SQL Injection'"
 SecRule ARGS|XML:/* "<script" "id:2,phase:2,t:none,t:htmlEntityDecode,t:lowercase,deny,status:400,msg:'Potential XSS Attack Detected'"
 SecRule RESPONSE_BODY "deface_pattern" "id:3,phase:4,t:none,t:lowercase,deny,status:403,msg:'Deface Attempt Detected'"
 SecRule REQUEST_HEADERS:User-Agent "libwww-perl" "id:4,phase:1,deny,status:403,msg:'Potential DDoS Tool: libwww-perl client'"
</IfModule>
```

Figure 5. Rules on VirtualHost Blocks

3.3. Intrusion Detection System (IDS) and Intrusion Prevention System (IPS) Configuration

Install snort and configure integration with waf located in snort.conf

```
 GNU nano 4.8                                                    snort.conf
#IDS
alert tcp any any -> any 80 (msg:"Possible Deface Attempt Detected"; content:"deface_pattern"; sid:100001; rev:1;)

alert tcp any any -> any 80 (msg:"Potential SQL Injection Detected"; content:"SELECT * FROM"; sid:100002; rev:1;)

alert tcp any any -> any 80 (msg:"Potential XSS Attack Detected"; content:"<script>"; sid:100003; rev:1;)

alert tcp any any -> any 80 (msg:"Potential DDoS Attack Detected"; flags:S; threshold: type both, track by_dst, count 100, seconds 10; sid:100004; rev:1;)

#IPS
alert tcp any any -> any 80 (msg:"SQL Injection Detected"; flow:to_server,established; content:"SELECT"; nocase; sid:100001; rev:1;)

alert tcp any any -> any 80 (msg:"Cross-Site Scripting (XSS) Attempt"; flow:to_server,established; content:"<script>"; nocase; sid:100002; rev:1;)

alert tcp any any -> any 22 (msg:"SSH Brute Force Attempt"; flow:to_server; flags:A+; threshold: type both, track by_src, count 5, seconds 60; sid:100003; rev:1;)

alert tcp any any -> $HOME_NET 21 (msg:"FTP Port Scan Detected"; flags:S; threshold: type both, track by_src, count 5, seconds 60; sid:100004; rev:1;)

alert icmp any any -> $HOME_NET any (msg:"ICMP Flood Detected"; dsize:0; threshold: type both, track by_src, count 50, seconds 10; sid:100005; rev:1;)
```

Figure 6. Snort integration with WAF

3.4. Rules Setting

Setting rules by pointing to Snort (IDS) first before passing to squid (as a reverse proxy) and WAF (Web Application Firewall). Using the management tool "iptables" to manage NAT (Network Address Translation) and routing.

```
root@webproxy:~# ^C
root@webproxy:~# ^C
root@webproxy:~# sudo iptables -A INPUT -p tcp --dport 80 -j REDIRECT --to-port 8888
root@webproxy:~# ^C
root@webproxy:~# sudo iptables -A INPUT -p tcp --dport 8888 -j REDIRECT --to-port 3128
root@webproxy:~# ^C
root@webproxy:~# sudo iptables -A INPUT -p tcp --dport 3128 -j REDIRECT --to-port 8080
root@webproxy:~# ^C
root@webproxy:~#
```

Figure 7. Rules Integration

This rule specifies that all packets arriving on port 80 (the default port for HTTP) are redirected to port 8888, where Snort is running. Snort is an intrusion detection system (IDS/IPS) that examines incoming traffic to detect potential security threats. It then redirects the traffic forwarded by Snort (running on port 8888) to Squid, running on port 3128. And forwards traffic routed through Squid (running on port 3128) to a WAF (such as ModSecurity) running on port 8080

3.5. Black Box Testing

To verify the security and robustness of the entire system, black box testing is performed with a predefined one with the aim of evaluating the external side of the system without detailed knowledge of its internal structure. Tests can be seen in the table.

Table 1. Black Box Testing

Features	Result Expectations	Result
Apache (Web Server)	accessed through a browser directs traffic to Apache as a web server	Valid
Squid (Reverse Proxy)	squid reverse proxy as the first layer before pointing to the Apache web server	Valid

Web Application Firewall (WAF)	WAF addresses attacks through the firewall in the form of defacers, SQL Injection, XSS, and DDOS.	Valid
Intrusion Detection System (IDS)	detects suspicious activity on the network. Sends an alert to the network or system administrator for further investigation.	Valid
Intrusion Prevention System (IPS)	detects suspicious activity, and takes action to prevent attacks, such as blocking traffic or changing system settings.	Valid

3.6. Integration Testing

This stage is to test the interaction between components that have been configured previously. Integration testing is done using Apache JMeter tools. Table 3 is the result of integration testing using Apache Jmeter.

Table 2. Integration Testing

Sample	**Before**			**After**		
Performance Test	Sample Time (ms)	Latency	Connect Time (ms)	Sample Time (ms)	Latency (ms)	Connect Time (ms)
10 User 1 Seconds 3 Loop	Min 174 Max 969	Min 114 Max 420	Min 173 Max 346	Min 71 Max 176	Min 71 Max 175	Min 73 Max 88
Stress Test	Sample Time (ms)	Latency	Connect Time (ms)	Sample Time (ms)	Latency	Connect Time (ms)
222 User 1 Seconds 2 Loop	Min 176 Max (969)	Min 110 Max 942	Min 177 Max 1128	Min 71 Max 401	Min 71 Max 210	Min 72 Max 114

Based on the test comparison table. Tests have shown that integrating squir reverse proxy as the first layer of the web server can significantly improve performance. This is evidenced by a 59% reduction in minimum sample time, a 58% reduction in maximum latency, and a 90% reduction in maximum connection time. These improvements show that integrating the Squid reverse proxy improves the efficiency and responsiveness of the web server, thereby improving the quality of service to users. The minimum sampling time before configuration was 174 ms and after configuration was 71 ms. The maximum latency before configuration was 420 ms and after configuration was 175 ms. The maximum connect time was 1128 ms before configuration and 114 ms after configuration. This performance improvement shows that integrating Squid reverse proxy allows the web server to process requests more quickly and efficiently, providing a better user experience.

3.7. Penetration Testing

This stage is to ascertain potential security or vulnerabilities in the system. Penetration testing is carried out with the aim of identifying and exploring security gaps.

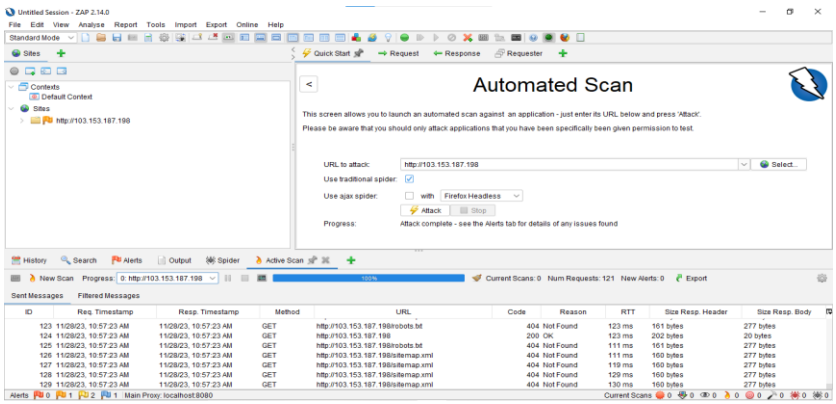


Figure 8. Penetration Testing

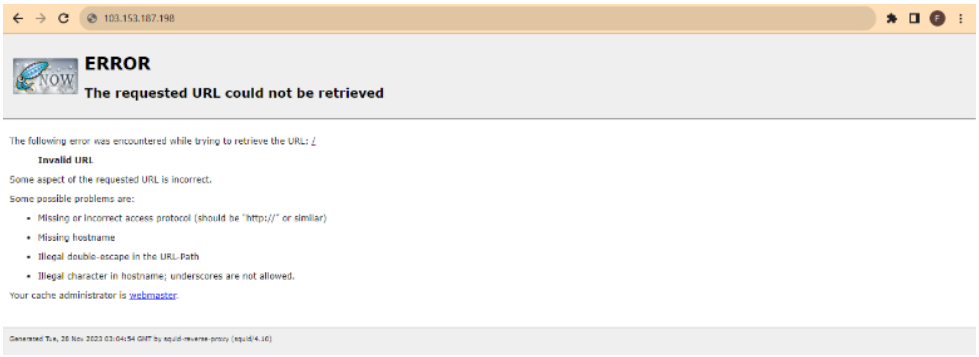


Figure 9. Squid Reverse Proxy Results After Penetration Testing

Penetration testing confirmed the success with the penetration results in Figure 7 showing that the configuration was successful where there were 0 critical vulnerabilities (red flags), 1 moderately severe vulnerability (orange flag), 2 vulnerabilities with the lowest level (yellow flag). from the penetration test, squid also prevented the attack by not forwarding the attack request to the main server and providing an error response to the attacker. Squid has proven effective in defending against attacks by blocking malicious requests and sending error messages to attackers. However, serious vulnerabilities must be addressed and continuous monitoring is essential. These upgrades have demonstrated a secure and efficient system, however continuous optimization is required to ensure maximum security.

4. Conclusion

After running a series of configurations, including implementing Squid as a reverse proxy, WAF (Web Application Firewall), and Snort as an IDS (Intrusion Detection System) and testing the effectiveness of the system, overall, the system successfully detected and prevented many common attacks such as SQL injection, XSS, and DDoS attacks. However, from the test still has a gap where there are still very critical vulnerabilities (high risk) which may cause great losses. if future researchers want to use this journal as a reference, it is hoped that they can develop even better security.

Acknowledgement

The researchers would like to express their deepest gratitude to Tadulako University, especially the Merdeka Center for Learning at Merdeka Campus (MBKM) and the Informatics Engineering Study Program. Thanks to the support of Tadulako University through the MBKM program, researchers can complete this research as well as possible. The opportunity to participate in the MBKM Independent Study program also provided the researchers with valuable experiences that increased their knowledge and insight in the field of computer science.

References

[1]. R. Riswandi, K. Kasim, and M. F. Raharjo, "Evaluasi Kinerja Web Server Apache menggunakan Protokol HTTP2," J. Tek. dan Sist. Inform., vol. 2, no. 1, pp. 19-31, 2020, doi: 10.36079/lamintang.jetas-0201.92.

[2]. Riska and H. Alamsyah, "Penerapan Sistem Keamanan Web Menggunakan Metode Web Application Firewall," J. Amplif. : J. Ilm. Bid. Tek. Elekt. dan Komput., vol. 11, no. 1, pp. 37-42, 2021, doi: 10.33369/jamplifier.v11i1.16683.

[3]. A. Susanto, "Pengaturan Keamanan Squid Proxy Pada Jaringan Lokal Dan Internet Menggunakan Openldap," 2022, vol. 2.

[4]. B. Wiguna, A. P. W., and R. Ananda, "Implementasi Web Application Firewall Dalam Mencegah Serangan SQL Injection Pada Website," Digital Zone, vol. 11, no. 2, pp. 245-256, 2020, doi: 10.31849/digitalzone.v11i2.4867.

[5]. T. A. F. Anugrah, S. Ikhwan, and A. G. J. Gusti, "Implementasi Intrusion Prevention System (IPS) Menggunakan Suricata Untuk Serangan SQL Injection," Techne, vol. 21, no. 2, pp. 199-210, 2022, doi: 10
[6]. Zain AR, Matin IMM. "Analisis Implementasi Modsecurity dan Reverse Proxy Untuk Pencegahan Serangan Keamanan DDoS pada Web Server." 2023;2(1).
[7]. Syani M. "IMPLEMENTASI INTRUSION DETECTION SYSTEM (IDS) MENGGUNAKAN SURICATA PADA LINUX DEBIAN 9 BERBASIS CLOUD VIRTUAL PRIVATE SERVERS (VPS)." JurnalInkofar. 2020;1(1). doi:10.46846/jurnalinkofar.v1i1.155
[8]. Suwanto R, Ruslianto I, Diponegoro M. "IMPLEMENTASI INTRUSION PREVENTION SYSTEM (IPS) MENGGUNAKAN SNORT DAN IPTABLE PADA MONITORING JARINGAN LOKAL BERBASIS WEBSITE." 2019;07(1).
[9]. Sutabri T. "Analisis Cyber Crime handling pada Aplikasi Web dengan WAF ModSecurity." 2023;16(1).
[10]. Ullah MU, Hassan A, Asif M, Farooq MS, Saleem M, Ullah U. "Intelligent Intrusion Detection System for Apache Web Server Empowered with Machine Learning Approaches." 2022;1.
[11]. Nasir A, Lordianto RL. "IMPLEMENTASI PROXY SERVER UNTUK OPTIMALISASI MANAJEMEN BANDWIDTH JARINGAN KOMPUTER PADA UNIVERSITAS XYZ." JTech. 2023;11(1):16-23. doi:10.30869/jtech.v11i1.1164
[12]. Muntaha MF, Trisnawan PH, Primananda R. "Implementasi Intrusion Prevention System (IPS) berbasis Athena untuk Mencegah Serangan DDoS pada Arsitektur Software-Defined Network (SDN).".
[13]. Muqorobin M, Hisyam Z, Mashuri M, Hanafi H, Setiyantara Y. "Implementasi Network Intrusion Detection System (NIDS) Dalam Sistem Keamanan Open Cloud Computing." MIBJ. 2019;17(2):1-9. doi:10.33489/mibj.v17i2.205
[14]. Universitas Methodist Indonesia, Gea A, Purba MJ, Putra AA, Jamaluddin J, Siringoringo R. "IMPLEMENTASI METODE ACCESS CONTROL LIST UNTUK MEMONITORING AKSES JARINGAN MENGGUNAKAN SQUID PROXY." jmika. 2022;6(1):79-84. doi:10.46880/jmika.Vol6No1.pp79-84
[15] R. A. Muzaki, O. C. Briliyant, M. A. Hasditama, and H. Ritchie, "Improving Security of Web-Based Application Using ModSecurity and Reverse Proxy in Web Application Firewall," in *2020 International Workshop on Big Data and Information Security (IWBIS)*, Depok, Indonesia: IEEE, Oct. 2020, pp. 85–90. doi: 10.1109/IWBIS50925.2020.9255601.
[16] R. A. Muzaki, Y. A. Wilar, K. Yuliawan, and A. A. Natsir, "ANALISIS KEAMANAN SISTEM MANAJEMEN INFORMASI RUMAH SAKIT UMUM DAERAH NABIRE," MAHESA: MALAHAYATI HEALTH STUDENT JOURNAL, vol. 3, no. 10, pp. 3365-3374, 2023, doi: 10.33024/mahesa.v3i10.11246.
[17] I. Stepheng, C. A. Sari, E. H. Rachmawanto, and F. O. Isinkaye, "A Combination of Vigenere Cipher and Advanced Encryption Standard for Image Security," ASSET: International Journal of Advanced Science and Engineering Technology, vol. 5, no. 3, pp. 0230305-01~0230305-08, Oct. 2023, doi: 10.26877/asset.v5i3.17150.
[18] A. S. Joel, F. Abdussalaam, dan Y. Yunengsih, "Tata Kelola Rekam Medis Berbasis Teknologi Informasi dalam Penanganan Kerahasiaan dan Keamanan Data Pasien dengan Metode Kriptografi," Jurnal Ilmiah Manajemen Informatika dan Komunikasi, vol. 4, no. 3, pp. 12-24, Sep. 2023, doi: 10.35870/jimik.v4i3.287.
[19] H. Amri, A. A. Haryada, K. Abdi, dan A. Ikhwan, "Manajemen Resiko Keamanan Aset Informasi Pada Puskesmas Pancur Batu Tuntungan," Jurnal Informatika dan Teknologi, vol. 3, no. 1, pp. 10-19, Mar. 2023.
[20] R. S. A. Gusni, Kraugusteeliana, dan I. W. W. Pradnyana, "Analisis Tata Kelola Keamanan Sistem Informasi Rumah Sakit XYZ Menggunakan Cobit 2019 (Studi Kasus pada Rumah Sakit XYZ)," Jurnal Informatika dan Teknologi, vol. 5, no. 1, pp. 6-11, Jun. 2021.

[21] J. Wijaya, A. Megafitri, K. Khotimah, dan R. Astriratma, "Analisis dan Manajemen Risiko Keamanan Informasi pada Rumah Sakit Menggunakan Metode Octave Allegro (Studi Kasus: Rumah Sakit Umum Daerah Cengkareng)," Jurnal Informatika dan Teknologi, vol. 2, no. 2, pp. 13-25, Sept. 2021.
[22] A. T. Kurniawan, B. S. WA, dan A. Nasiri, "Analisis Architecture Teknologi Menggunakan SABSA untuk Meningkatkan Keamanan di Rumah Sakit Queen Latifa," Jurnal Informatika dan Teknologi, vol. 3, no. 2, pp. 6-11, Des. 2022.

Advance Sustainable Science, Engineering and Technology (ASSET)
Vol. 6, No.2, April 2024, pp. 02402010-01 ~ 02402010-08
ISSN: 2715-4211 DOI: https://doi.org/10.26877/asset.v6i2.18262

Implementing Blockchain For Publishing and Verifying Digital Certificates On EduTech

Akwan Maroso*, Dwi Shinta Angreni, Rizka Ardiansyah, Kadek Agus Dwiwijaya

Information Technology Department, Faculty of Engineering, Tadulako University, Jl. Soekarno Hatta No KM 9 Palu 94148, Central Sulawesi, Indonesia

*akwanmaroso@gmail.com

This study investigates the application of blockchain technology in enhancing the security and authenticity of digital certificates. Addressing key challenges such as fraud and the lack of a standardized verification process, the paper proposes a comprehensive framework aimed at fortifying the integrity of digital credentials. This framework is the utilization of blockchain as a distributed ledger, serving as a tamper-proof repository for recording certification transactions. Through this decentralized ledger, each certification issuance and verification action is securely recorded, enhancing trust and transparency in the certification process. The methodology includes the integration of a decentralized ledger for immutable record-keeping and implementation of smart contracts for automated authenticity checks, and the use of cryptographic measures to ensure data security. This approach promises significant implications for various sectors reliant on credential verification, advocating for a broader adoption of blockchain in digital certificates systems.

Keywords: Blockchain, Digital Certificates, Smart Contract, Decentralization, Data Security.

(Received 2024-01-25, Accepted 2024-02-20, Available Online by 2024-03-08)

1. Introduction

For almost 20 years, digital certificates have been a key aspect of Internet security [1]. The need for reliable verification of credentials has become increasingly important. Digital certificates, serving as electronic documents to certify the ownership of a public key, have emerged as a critical tool in various fields ranging from education to IT security. They are used to validate the authenticity of entities and secure communication over the internet [2].

Traditional methods of issuing and verifying digital certificates face several challenges. These include susceptibility to fraud, lack of transparency, and inefficiencies in verification processes. The centralized nature of traditional systems often leads to a single point of failure, making them vulnerable to cyber-attacks. Additionally, the process of verifying certificates can be cumbersome and time-consuming, requiring reliance on the issuing authority for validation [3].

Blockchain technology, best known for underpinning cryptocurrencies like Bitcoin, offers a decentralized and tamper-evident ledger system. Its inherent characteristics—decentralization, immutability, and transparency—make it an appealing solution to address the shortcomings of

traditional digital certificate systems [4].

Implementing blockchain technology in the issuance and verification of digital certificates introduces a paradigm shift. Blockchain's decentralized nature allows for the creation of a distributed ledger where each certificate can be securely and transparently recorded. This approach reduces the risk of fraud and unauthorized alterations [5].

Despite the potential benefits of the blockchain system, the application of blockchain in digital certificates is still in its nascent stage, with practical challenges and limitations not fully explored. This research aims to fill this gap by implementing and analyzing blockchain technology in the context of digital certificate issuance and verification. The study's findings are expected to provide valuable insights and contribute to the advancement of knowledge in this evolving field.

To guide readers through the content of this paper, the following is the structure that will be followed. The next section will delve into the Methods, discussing the research type employed and the system development undertaken. Subsequently, the paper will move on to the Results and Discussion section, covering planning, analysis, system design, prototype, implementation, and system testing. The paper will conclude with a summary and suggestive recommendations for future developments in this field.

2. Methods

The research methodology is illustrated in Figure 1, outlining the flow of the study. This process begins with problem identification, followed by data collection, system design, system implementation, system testing, and ultimately concluding with the findings.

Figure 1. Research Methodology

2.1. Identification Problem

In this section, we'll examine the existing process of publishing and verifying digital certificates, identifying potential issues along the way.

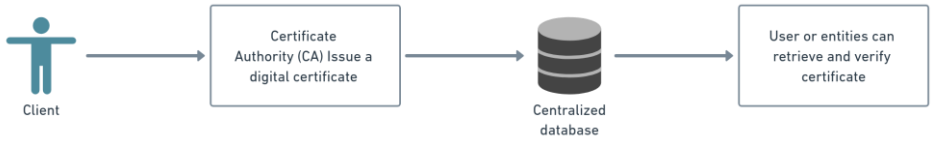


Figure 2. Current Publish Certificate System

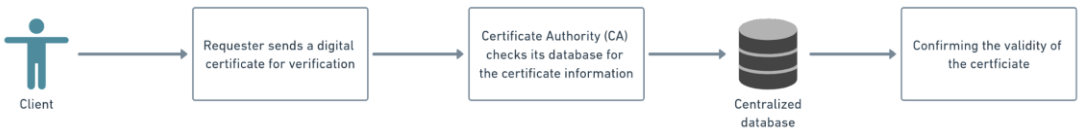


Figure 3. Current Verifying Certificate System

Figure 2 illustrates the process flow for publishing a certificate. The Certificate Authority (CA) stores data in a centralized database, such as MySQL or PostgreSQL [6], and makes the digital certificate accessible to entities. Figure 3 outlines the certificate verification process, where the client or requester sends a validation request. The CA then checks the centralized database for the provided information and confirms the validity of the certificate. The problem arises from using a centralized database to store certificate data. This creates single point failure and vulnerability to security breaches or data

manipulation [7]. Additionally, relying solely on the CA to verify certificate introduce a level of trust that may be exploited.

2.2. Data Collection

This phase focuses on gathering the necessary data to develop the system. Using quantitative techniques to generate large numbers of observations, emphasizing the necessity of large sample sizes for sophisticated quantitative analyses [8]. Data collected are associated with the process of publication and verification of digital certificates within the EduTech community Kode Karya Palu (Hammercode) are addressed. The research involved compiling recent digital certificates that have been published in the community.

2.3. System Design

In this stage, we defined how we can integrate the process of publish and verifying the certificate. The application was programmed under the Ethereum platform. Ethereum is major blockchain-based platform for smart contracts with complete programs that are executed in a decentralized network and usually manipulate digital units of value. A peer-to-peer network of mutually distrusting nodes maintains a common view of the global state and executes code upon request. The stated is stored in a blockchain secured by a proof-of-work consensus mechanism similar to that in Bitcoin [9-11]. Smart contracts will act as stateful decentralized applications that run on EVM implementations to enforce contract instructions [12]. In this system, we will use the RESTful API [13] as the protocol communication between client app and server app.

In Figure 4 we can see the process of publishing a certificate in the initial step, certificate details are input into the system and will trigger the smart contract within the EVM, utilizing the provided information. The smart contract processes the request according to the pre-defined contract, and upon successful fulfillment, stores the data as a transaction within the blockchain. Each process that creating transactions on the blockchain will be need a cost (gas) to distribute the data across each node [14][15]. Upon completion of all processes, the system generates a random hash along with a QR code to serve as an identifier for the certificate stored in the blockchain.

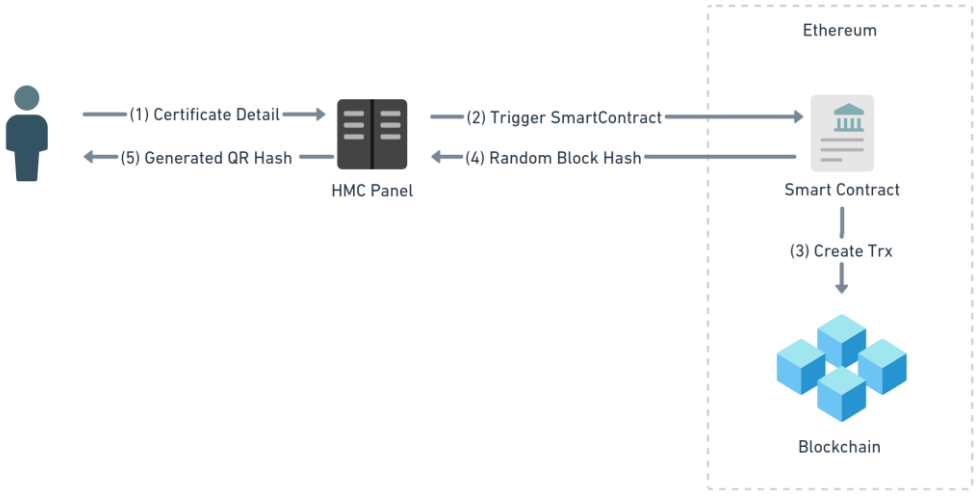


Figure 4. Publish Certificate

On Figure 5 illustrates the process of verifying a certificate. In this workflow, the client sends a hashed QR code containing the certificate identifier. The system acts as a proxy, triggering the smart contract equipped with a function to locate the certificate based on the identifier. Since the process involves only reading data from the blockchain network, there are no additional costs incurred [14]. If the data is found, it is displayed to the client.

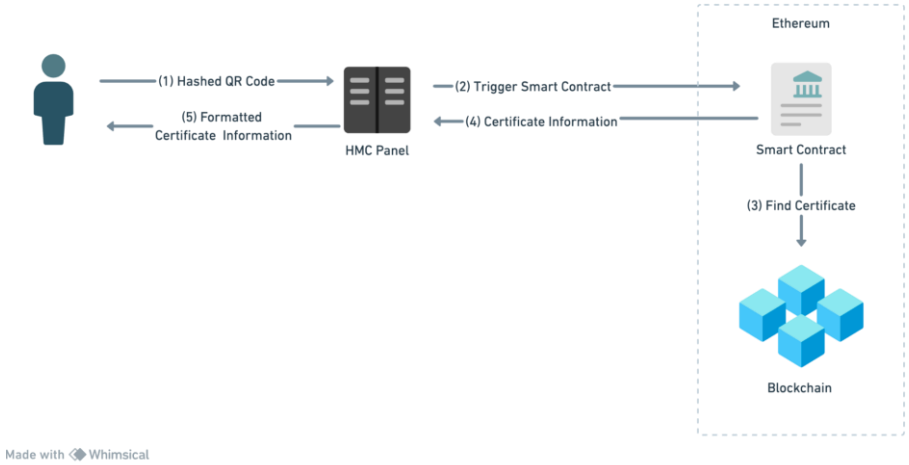


Figure 5. Verify Certificate

2.4. Implementation

The system has separate backend and frontend components. The backend is built using the Go programming language. Go is expressive, concise, clean, and efficient. Its concurrency mechanisms make it easy to write programs that get the most out of multicore and networked machines, while its novel type system enables flexible and modular program construction [16], while the frontend development incorporates JavaScript, utilizing NextJS as the framework. Additionally, the system's interaction with the Ethereum network is facilitated through smart contracts written in Solidity[17].

The client-side interface allows for two methods of inputting the certificate ID. Firstly, users can manually enter the certificate ID and initiate verification; the system will then validate the certificate. Alternatively, users can scan the QR code on the certificate, which automatically extracts the ID. This implementation is illustrated in Figure 6.

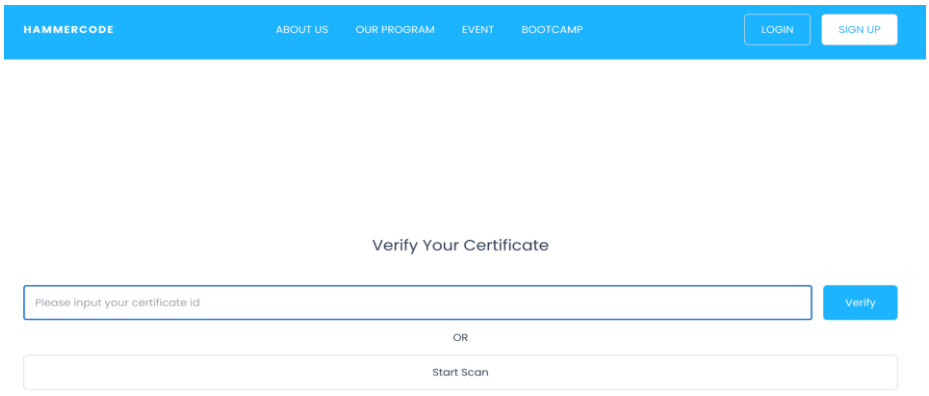


Figure 6. Validation Page

When a value is inputted as shown in Figure 6, it undergoes verification within the blockchain network. Should the data be found on the network, the details of the certificate will be presented on the system. The outcomes of this verification process are illustrated in Figure 7.

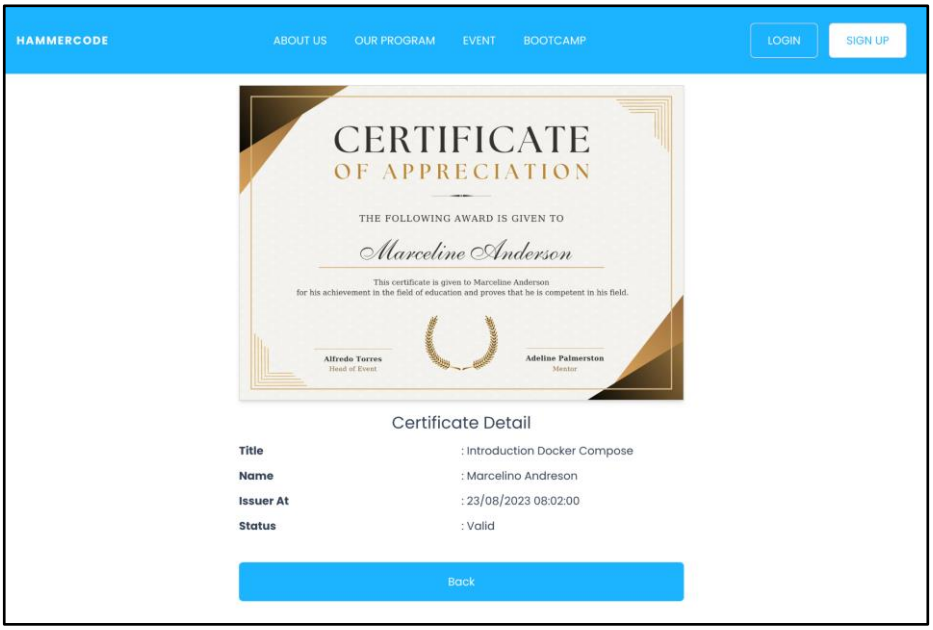


Figure 7. Certificate Validation Result

2.5. Testing Phase

In this study, the performance evaluation of the application hinges on analyzing key metrics. These metrics encompass the calculation of transaction throughput (TTP) and verification latency (VL). By focusing on these parameters, we aim to gain insights into the efficiency and responsiveness of the application.

1. Transaction Throughput (TPP)
   This metric reflects of the number of certificate verification transactions the system can handle per unit of time[18], indicating scalability and performance efficiency.

$$TTP = \frac{Total\ Number\ of\ Transaction\ Processed}{Total\ Time\ Taken}$$

2. Verification Latency
   Average time taken to verify a single certificate, this indicate the responsiveness the system [18].

$$VL = \frac{Total\ Time\ Taken\ to\ Verify\ N\ Certificate}{N}$$

3. Results and Discussion

3.1. Result

In our assessment of the blockchain application developed for issuing and verifying digital certificates, we noted significant achievements in several metrics, along with areas that require enhancement, particularly addressing the issue of a single point of failure in the database. The outcomes of our performance analysis, focusing on Transaction Throughput(TTP) and Verification Latency(VL), are presented below.

$$TTP = \frac{135}{4}$$
$$TTP = 33\ transaction\ per\ second$$

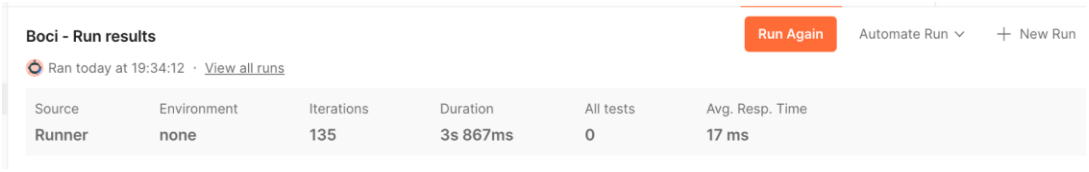


Figure 8. Load Test Publish Certificate

$$_{V}L = \frac{9}{135}$$

$$VL = 60ms\ per\ certificate$$

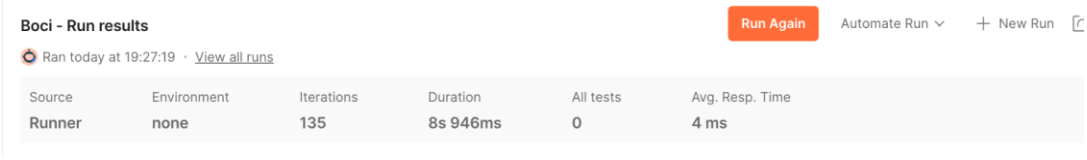


Figure 9. Load Test Verify Certificate

The detailed examination of our load testing, conducted using Postman as the load testing client [19] on Figure 8 and Figure 9, involved a scenario with 135 certificates undergoing the process of publication and verification. This thorough evaluation demonstrated a Transaction Throughput (TTP) of 33 transactions per second. Concurrently, the Verification Latency (VL) for each transaction was recorded at 60 milliseconds. These results underscore a significant opportunity for enhancing the system's speed, particularly in optimizing the verification process to achieve a more efficient transaction handling rate.

3.2. Blackbox Testing

This stage will test all process that are already develop. The testing process gonna use Black Box testing method also know a functional testing or input output driven testing, it means on the process testing will be focus on the input and output of the system [20]. The following are the results of system testing using the Blackbox method can be seen in Table 1.

Table 1. System Testing Results

Test Case	Description	Results
Deploy smart contract	Test the deployment of the smart contract on the Ethereum network.	Success
Verify Certificate	Validate if the inputted certificate ID is valid or not.	Success
Publish Certificate	Store certificate details on the blockchain network.	Success
Invalid Certificate Check	Verify system's response to an invalid certificate ID.	Success
Certificate Data Retrieval	Retrieve the details of a stored certificate.	Success
Smart Contract Update	Test updating the smart contract code on the Ethereum network.	Success

4. Conclusion

Based on the system testing result and observation this research about adopting blockchain to process publishing and verifying certificate digital with blockchain has been successfully implemented. This adoption make the each process publish and verify certificate more transparent and more secure. This approach not only enhances on factor integrity and security but also expand solution to the limitations associated with traditional systems. The results of the system testing provided us visual evidence of the system accelerated speed. These result offer significantly improved processing times, showcasing the efficiency gains achieved through the adoption of this technology

One of the critical advantages addressed in this study is blockchain's ability to mitigate the risks associated with single points of failure in centralized database systems. In traditional systems, the centralization of data storage and management creates vulnerabilities, making the entire system susceptible to data breaches, tampering, and downtime. Future research could focus on further exploring the scalability of blockchain solutions on the digital certificates, examining the integration with existing educational and professional credentialing systems, and developing more advanced cryptographic methods to enhance privacy and security even further. Additionally, investigating mechanisms to address potential scalability challenges and ensuring the interoperability of blockchain-based certification systems across various sectors and platforms would be invaluable.

References

[1] N. Leavitt, "Internet Security under Attack: The Undermining of Digital Certificates," in Computer, vol. 44, no. 12, pp. 17-20, Dec. 2011, doi: 10.1109/MC.2011.367.

[2] Maulani, G., Gunawan, G., Leli, L., Ayu Nabila, E., & Yestina Sari, W. (2021). Digital Certificate Authority with Blockchain Cybersecurity in Education. *International Journal of Cyber and IT Service Management*, *1*(1), 136–150. https://doi.org/10.34306/ijcitsm.v1i1.40

[3] Jirgensons, M., & Kapenieks, J. (2018). Blockchain and the Future of Digital Learning Credential Assessment and Management. *Journal of Teacher Education for Sustainability*, *20*(1), 145–156. https://doi.org/10.2478/jtes-2018-0009

[4] S. Nakamoto, “Bitcoin: A Peer-to-Peer Electronic Cash System,” Www.Bitcoin.Org, p. 9, 2008.

[5] T. T. Huynh, T. Tru Huynh, D. K. Pham and A. Khoa Ngo, "Issuing and Verifying Digital Certificates with Blockchain," 2018 International Conference on Advanced Technologies for Communications (ATC), Ho Chi Minh City, Vietnam, 2018, pp. 332-336, doi: 10.1109/ATC.2018.8587428.

[6] Poljak, R., Poščić, P., & Jakšić, D. (2017). Comparative analysis of the selected relational database management systems. 2017 40th International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO), 1496-1500. https://doi.org/10.23919/MIPRO.2017.7973658.

[7] Jung, M. Y., & Jang, J. W. (2017). Data management and searching system and method to provide increased security for IoT platform. Information and Communication Technology Convergence, 2017-December, 873–878. https://doi.org/10.1109/ICTC.2017.8190803.

[8] Ragin, C.C. (1999). The distinctiveness of case-oriented research. Health services research, 34 5 Pt 2, 1137-51.

[9] Tikhomirov, S. (2018). Ethereum: State of Knowledge and Research Perspectives. In: Imine, A., Fernandez, J., Marion, JY., Logrippo, L., Garcia-Alfaro, J. (eds) Foundations and Practice of Security. FPS 2017. Lecture Notes in Computer Science(), vol 10723. Springer, Cham. https://doi.org/10.1007/978-3-319-75650-9_14

[10] D. Vujičić, D. Jagodić and S. Ranđić, "Blockchain technology, bitcoin, and Ethereum: A brief overview," 2018 17th International Symposium INFOTEH-JAHORINA (INFOTEH), East Sarajevo, Bosnia and Herzegovina, 2018, pp. 1-6, doi: 10.1109/INFOTEH.2018.8345547.

[11] Oliva, G.A., Hassan, A.E. & Jiang, Z.M.(. An exploratory study of smart contracts in the Ethereum blockchain platform. Empir Software Eng 25, 1864–1904 (2020). https://doi.org/10.1007/s10664-019-09796-5

[12] Chan, W., & Olmsted, A. (2017). Ethereum transaction graph analysis. 2017 12th International Conference for Internet Technology and Secured Transactions (ICITST), 498–500. https://doi.org/10.23919/ICITST.2017.8356459
[13] Ehsan, A., Abuhaliqa, M. A. M. E., Catal, C., & Mishra, D. (2022). RESTful API Testing Methodologies: Rationale, Challenges, and Solution Directions. In Applied Sciences (Switzerland) (Vol. 12, Issue 9). MDPI. https://doi.org/10.3390/app12094369
[14] L. Marchesi, M. Marchesi, G. Destefanis, G. Barabino and D. Tigano, "Design Patterns for Gas Optimization in Ethereum," 2020 IEEE International Workshop on Blockchain Oriented Software Engineering (IWBOSE), London, ON, Canada, 2020, pp. 9-15, doi: 10.1109/IWBOSE50093.2020.9050163.
[15] B. Severin, M. Hesenius, F. Blum, M. Hettmer and V. Gruhn, "Smart Money Wasting: Analyzing Gas Cost Drivers of Ethereum Smart Contracts," 2022 IEEE International Conference on Software Maintenance and Evolution (ICSME), Limassol, Cyprus, 2022, pp. 293-304, doi: 10.1109/ICSME55016.2022.00034.
[16] Russ Cox, Robert Griesemer, Rob Pike, Ian Lance Taylor, and Ken Thompson. 2022. The Go programming language and environment. Commun. ACM 65, 5 (May 2022), 70–78. https://doi.org/10.1145/3488716.
[17] Analyzing the Impact of Next.JS on Site Performance and SEO. (2023). International Journal of Computer Applications Technology and Research. https://doi.org/10.7753/ijcatr1210.1004.
[18] Yadav, J., & Shevkar, R. (2021). Performance-Based Analysis of Blockchain Scalability Metric. Tehnicki Glasnik, 15(1), 133–142. https://doi.org/10.31803/TG-20210205103310.
[19] Andrianto, L. D., & Suyatno, D. F. (2024). Analisis Performa Load Testing Antara Mysql Dan Nosql Mongodb Pada RestAPI Nodejs Menggunakan Postman. Journal of Emerging Information System and Business Intelligence (JEISBI), 5(1), 18–26. https://ejournal.unesa.ac.id/index.php/JEISBI/article/view/58157.
[20] Syarif, M., & Pratama, E. B. (2021). ANALISIS METODE PENGUJIAN PERANGKAT LUNAK BLACKBOX TESTING DAN PEMODELAN DIAGRAM UML PADA APLIKASI VETERINARY SERVICES YANG DIKEMBANGKAN DENGAN MODEL WATERFALL. *Jurnal Teknik Informatika Kaputama (JTIK)*, *5*(2).

Advance Sustainable Science, Engineering and Technology (ASSET)
Vol. 6, No.2, April 2024, pp. 02402011-01 ~ 02402011-09
ISSN: 2715-4211 DOI: https://doi.org/10.26877/asset.v6i2.18308

Wina Sentosa Bottled Water Distribution System Using Web-Based Distribution Requirement Planning and Trend Moment Algorithms

Jonathan Albert Uswary[1], Hajra Rasmita Ngemba[2*], Syaiful Hendra[3], Anita Ahmad Kasim[4]

[1,3,4]Information Technology Department, Faculty of Engineering, Tadulako University, Jl. Soekarno Hatta No KM 9 Palu 94148, Central Sulawesi, Indonesia
[2]Information System Department, Faculty of Engineering, Tadulako University, Jl. Soekarno Hatta No KM 9 Palu 94148, Central Sulawesi, Indonesia
*hajra.rasmita@gmail.com

Abstract. This research focuses on the problem of PT Anugrah Wina Sentosa, which is a producer of bottled drinking water in Central Sulawesi. The company faces challenges in organizing and improving the efficiency of the distribution of bottled drinking water products. Problems include distribution management that has not been optimized and distribution arrangements at various outlets that have not been mapped. Based on this problem, researcher develop a web-based distribution system with the Distribution Requirement Planning (DRP) algorithm and Trend Moment Algorithm to see the results of sales and distribution predictions. The designed application can carry out data processing, Distribution Requirement Planning processes, Trend moment processes, MAPE calculations, and sales predictions. The application development method uses the Waterfall method. The test results show that the system can manage input, edit, and delete data and run DRP calculations as a whole or per outlet. With an error value of 1.71%, the trend moment forecasting system proved to be very accurate in forecasting sales of drinking water products. Thus, the implementation of a web-based distribution system can improve production efficiency and facilitate stock management and distribution management at PT Anugrah Wina Sentosa. This research has several limitations that need to be considered, namely limited scope, limited data, limited generalization, limited affordability, and limited time.

Keywords: Distribution System, Website, Distribution Requirement Planning (DRP), Trend Moment, Mean Absolute Percentage Error (MAPE)

(Received 2024-01-31, Accepted 2024-02-13, Available Online by 2024-03-02)

1. Introduction

Indonesia is the fifth largest bottled water market in the world. Many bottled drinking water companies have spread in every region in Indonesia [1]. The growth of the food and beverage industry in Indonesia contributed greatly to national investment by contributing up to IDR 56.60 trillion in 2018, and the food and beverage industry grew by 7.91%, exceeding national economic growth by 5.17% [2]. People in Indonesia make bottled drinking water a daily product; the National Socio-economic Survey (SUSENAS) conducted by the Central Statistics Agency (BPS) in March 2023 showed that 40.64% of households in Indonesia chose to use bottled drinking water / refilled water daily drinking water. Along

with the increasing population in Indonesia, the need for bottled drinking water will also increase [3]. Efforts to increase product demand and a reliable distribution system are needed to meet product demand. Distribution is the process of distributing goods or services from producers to consumers [4],[5]. The distribution system is one of the important factors for companies to deliver products appropriately to customers [6]. This is what was done by PT Anugrah Wina Sentosa.

PT Anugrah Wina Sentosa is one of the bottled drinking water producers in Central Sulawesi with a product called WN. WN is a local brand of bottled drinking water with international standards. WN has several size variants, namely, 19L gallons, 330ml and 600ml bottles, and 220ml and 250ml cups. The 19L gallon WN has two product variations, namely, original WN gallons and WN gold gallons. Distribution areas include West Palu, East Palu, North Palu, South Palu, and Tatanga, with a total of more than 1,000 outlets. In the West Palu area, WN gallons are distributed to 179 different outlets. The large demand for products caused PT Anugrah Wina Sentosa to experience several obstacles in organizing and increasing the efficiency of the distribution of drinking water products in WN packaging. These problems include aspects such as distribution management that is not yet optimal, increasing the number of outlets that can be reached, and distribution arrangements in various areas of Palu City that have not yet been mapped. If this problem continues it can cause inventory shortages, increased distribution costs, and an inability to meet consumer demand.

The implementation of a web-based distribution system can provide several benefits for PT Anugrah Wina Sentosa, especially in terms of operational efficiency and responsiveness to market changes. The application design will use the Distribution Requirement Planning algorithm and the Trend Moment Algorithm. Distribution Requirement Planning (DRP) is an algorithm for managing product inventory and distribution so that it can carry out distribution design to meet consumer demand at the right time and amount [7]. The author chose the DRP algorithm because the problems faced by PT Anugrah Wina Sentosa were by the objectives of this algorithm, namely inventory management and product distribution planning to meet customer demand. Trend moment is a method for forecasting using visualization of trend lines in sales [8]. The trend moment method in forecasting has several advantages compared to other methods, especially for forecasting data that shows a significant trend or pattern of increase or decrease.This research seeks to explore the potential benefits of a web-based distribution system in increasing product production efficiency. The main objective is to assist the stock management process and optimize product distribution strategies, to meet customer demand, thereby contributing to overall system operational improvements.

2. Methods

2.1. System Flowchart

A system flowchart is a flowchart that shows the sequence chart of a system's operating process [9]. The system to be created is a WN gallon water distribution system using the Distribution Requirement Planning algorithm and the web-based Trend Moment Algorithm. The system starts from logging in as an admin on the website page. When the login is successful, the admin can input sales data. After inputting the data, several data processing processes will be carried out. Namely, the DRP process, the Trend Moment process, and finally, the MAPE test from the Trend Moment process results so that it can produce forecasting for the next month's sales and see the error accuracy level of MAPE testing. The system flow can be seen in Figure 1.

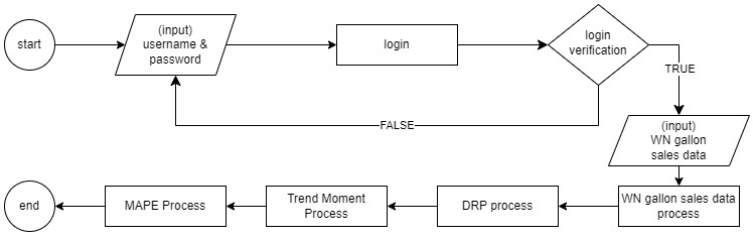


Figure 1. System flowchart

2.2. Dataset

Datasets are collections of objects and properties or characteristics of an object itself [10]. The dataset used in this study is WN gallon water sales data from April 2023 to October 16, 2023, in the West Palu region, Central Sulawesi, with a total of 179 outlets and total sales of 4267. The dataset used can be added or changed to adjust to the latest sales situation of PT Anugrah Wina Sentosa so that the application can be used continuously. The author has been permitted to use the sales data in this research.

	TANGGAL TRAX	NO. NOTA	TOP	NAMA CUSTOMER	ALAMAT	REA	BARANG YANG DIPESAN	QTY	HARGA SATUAN	TOTAL
2	01 April 2023	CASH	0	AKBAR/		BARAT	WN Galon Gold	3	Rp 14.000	Rp 42.000
3	01 April 2023	CASH	0	ALFI		BARAT	WN Galon Gold	3	Rp 14.000	Rp 42.000
4	01 April 2023	CASH	0	ALFI		BARAT	WN Galon Original	8	Rp 11.000	Rp 88.000
5	01 April 2023	CASH	0	ANDIKA/..		BARAT	WN Galon Gold	11	Rp 14.000	Rp 154.000
6	01 April 2023	CASH	0	ANDIKA/..		BARAT	WN Galon Original	2	Rp 11.000	Rp 22.000
7	01 April 2023	CASH	0	DEPOT, ANISA		BARAT	WN Galon Gold	8	Rp 14.000	Rp 112.000
8	01 April 2023	CASH	0	DELTA		BARAT	WN Galon Gold	12	Rp 14.000	Rp 168.000
9	01 April 2023	CASH	0	FAMILY 2		BARAT	WN Galon Gold	25	Rp 14.000	Rp 350.000
10	01 April 2023	CASH	0	HASNA/..		BARAT	WN Galon Gold	9	Rp 14.000	Rp 126.000
11	01 April 2023	CASH	0	IDA		BARAT	WN Galon Gold	2	Rp 14.000	Rp 28.000
12	01 April 2023	CASH	0	INDRI, KIOS		BARAT	WN Galon Gold	14	Rp 14.000	Rp 196.000
13	01 April 2023	CASH	0	INDRI, KIOS		BARAT	WN Galon Original	5	Rp 11.000	Rp 55.000
14	01 April 2023	CASH	0	KIKI II		BARAT	WN Galon Gold	11	Rp 14.000	Rp 154.000
15	01 April 2023	CASH	0	MARIAM		BARAT	WN Galon Gold	20	Rp 14.000	Rp 280.000
16	01 April 2023	CASH	0	NIA		BARAT	WN Galon Gold	7	Rp 14.000	Rp 98.000
17	01 April 2023	CASH	0	PRIMA,		BARAT	WN Galon Gold	4	Rp 14.000	Rp 56.000
18	01 April 2023	CASH	0	PRIMA PEMBAHARU		BARAT	WN Galon Gold	11	Rp 14.000	Rp 154.000
19	01 April 2023	CASH	0	RAHMAT		BARAT	WN Galon Gold	7	Rp 14.000	Rp 98.000
20	01 April 2023	CASH	0	REZA		BARAT	WN Galon Gold	4	Rp 14.000	Rp 56.000
4263	16 October 2023	CASH	0	SAL, TOKO		BARAT	WN Galon Gold	24	Rp 14.000	Rp 336.000
4264	16 October 2023	CASH	0	SAL, TOKO		BARAT	WN Galon Original	9	Rp 11.000	Rp 99.000
4265	16 October 2023	CASH	0	SAMIRAW		BARAT	WN Galon Gold	13	Rp 14.000	Rp 182.000
4266	16 October 2023	CASH	0	KIOS		BARAT	WN Galon Gold	19	Rp 14.000	Rp 266.000
4267	16 October 2023	CASH	0	UDIN/		BARAT	WN Galon Gold	9	Rp 14.000	Rp 126.000
4268	16 October 2023	CASH	0	ZAMRUD /		BARAT	WN Galon Gold	4	Rp 14.000	Rp 56.000

Figure 2. Sales dataset from April 01, 2023-October 16, 2023

2.3. Distribution Requirement Planning

The Distribution Requirement Planning (DRP) algorithm is an algorithm for handling product inventory stocks in a company to carry out the distribution [7]. The DRP algorithm can manage product availability and determine distribution planning so as to improve the company's sales performance in meeting consumer demand at the right time with the right amount [5]. The terms used in this algorithm are Safety Stock (SS), Period, Lot Size (LS), Lead Time (LT), Past Due (PD), Gross Requirement (GR), Projected on Hand (POH), Net Requirement (NR), Planned Order Receipts and Planned Order Release. The following equation calculates the value of Projected on Hand:

$$POH = POH_{(t-1)} - GR + POR \quad (1)$$

Description:
POH = projected on hand
$POH_{(t-1)}$ = previous period's projected on hand
GR = gross requirement

The following equation calculates the net requirement value.

$$NR = GR + SS - POH_{(t-1)} \quad (2)$$

Description:
NR = net requirement
GR = gross requirement
SS = safety stock
$POH_{(t-1)}$ = previous period's projected on hand

2.4. Trend Moment

The Trend Moment method is used to find the trend line using statistical and mathematical calculations to see the intersection between the company's historical data [11]-[13]. The trend moment algorithm is chosen due to its excellence in capturing trends and fluctuations in data, which are essential characteristics in forecasting within distribution systems. Within distribution systems, the trend moment algorithm contributes to improvement by providing more accurate forecasts of future demand. To do forecasting, this method uses the equation:

$$Y = a + bX$$

Description:
Y = trend value
a = constant number
b = slope or trend line coefficient
X = time index (0,1,2,....n)

To find the values of a and b, the following elimination function is used:

$$\sum y = (n \times a) + (b \times \sum x) \quad (3)$$

$$\sum xy = (a \times \sum x) + (b \times \sum x^2) \quad (4)$$

Description:
$\sum y$ = number of sales data
$\sum x$ = number of time periods
$\sum xy$ = number of sales data times the time period
n = amount of data

2.5. Mean Absolute Percentage Error

Mean Absolute Percentage Error (MAPE) is a matrix that aims to calculate the error value or accuracy of forecasting performed by the Trend Moment method by comparing the specified test data and training data. The results of these calculations can illustrate the accuracy of the prediction results of the trend moment method using training data. [14]. The MAPE calculation formula is:

$$\text{MAPE} = \frac{1}{n}\sum \left|\frac{y_t - y'_t}{y_t}\right| \times 100 \quad (5)$$

Description:
n = sample size
y_t = actual value at time t
y'_t = predicted or forecasting value at time t

From the MAPE calculation formula, the actual value is reduced by the predicted value and then divided by the actual value, which is made into an absolute value and multiplied by 100. The grouping of MAPE value categories can be seen in Table 1.

Table 1. MAPE grouping

MAPE Value	Category
<10%	Very accurate
10-20%	Good
20-50%	Reasonable
>50%	Inaccurate

3. Results and Discussion

3.1. System Testing Result

After testing, the following are the results of processing WN gallon water sales data from April 01 to October 16, 2023. The designed system provides good results; the existing features can function properly. On the dashboard page, it can be seen that there is total revenue, total sales of 4267, 179 outlets, two types of products, namely WN gold and original gallons, and a graph of revenue every month for a period of 1 year. The revenue graph in October looks very low because the dataset used is only until October 16. The dashboard display can be seen in Figure 3.

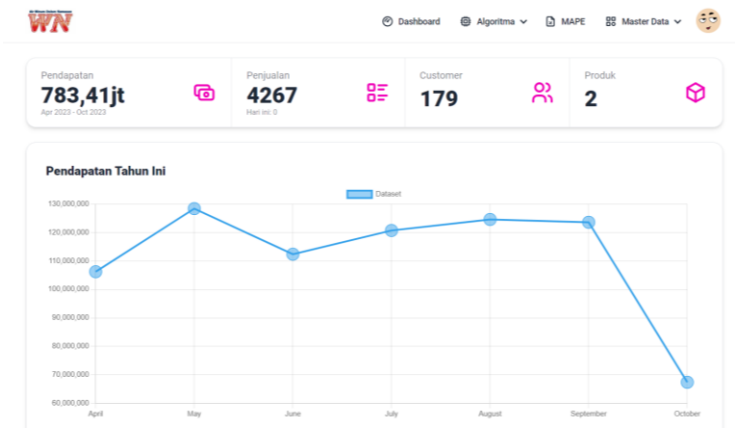


Figure 3. Dashboard

On the DRP algorithm page, the system automatically calculates the DRP of each outlet in the dataset. The example used is the Ahsan outlet. The DRP calculation of Ahsan outlet can be seen in Figure 4. DRP calculation starts with determining safety stock. Because the product is ordered by the customer, the safety stock is automatically not used, so the safety stock is equal to 0. The period in this calculation is based on sales per month. Then, calculate the past due period. Past due is the initial stage of the calculation. In the past due period, the number of products in the Ahsan outlet is unknown, so the projected number on hand is 0. In the April period, the customer gross requirement was 234, so the initial release order plan needed 234. When PD, the value of plan order release is 234, but in April, the value of plan order receipts is 234. Next, the net requirement calculation is carried out in the April period.

$$NR = GR + SS - POH_{(t-1)} \qquad NR = 234 + 0 - 0 \qquad NR = 234$$

Then, the NR value in April is 234. After completing the calculation in the April period, do the calculation again in the next period. Then the calculation results are shown in Figure 4.

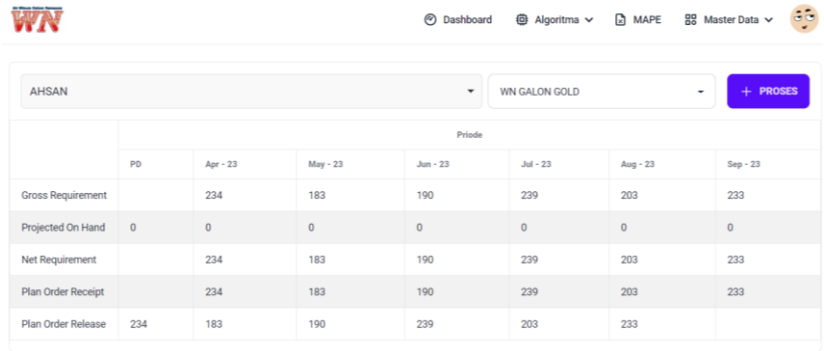


Figure 4. DRP Algorithm Process at Ahsan Outlet

On the trend moment algorithm page, the system automatically calculates the total sales for sales prediction in the next period based on the dataset used, can be seen in figure 5. The first three months are used, namely April, May, and June. The y value is the sales value. The y value is obtained based on the sales dataset used. The x value is the time in the sales period starting from 0. Furthermore, the calculation starts from the first period, namely April. The x value^2 is 0 because in the first period the x value is equal to 0. Then, the x*y value is 0 because, in the first period, the x value is equal to 0. Continue the calculation until the third period. After getting the value of each period, the next step is to create an equation using the formula $\sum y = (n \times a) + (b \times \sum x)$ and $\sum xy = (a \times \sum x) + (b \times \sum x^2)$.

$$\sum y = (n \times a) + (b \times \sum x) \qquad 22270 = (3 \times a) + (b \times 3) \qquad 22270 = 3a + 3b \qquad (1)$$

$$\sum xy = (a \times \sum x) + (b \times \sum x^2) \quad 22526 = (a \times 3) + (b \times 5) \quad 22526 = 3a + 5b \tag{2}$$

After getting the two equations above, elimination is done to find the value of b, and substitution is done to find the value of a.

$$\begin{array}{r} 22270 = 3a + 3b \\ \underline{22526 = 3a + 5b} \\ -256 = -2b \\ b = -\frac{256}{-2} \\ b = 128 \end{array}$$

$$\begin{aligned} 22270 = 3a + 3b22270 \\ &= 3a + 3 \times 12822270 \\ &= 3a + 3843a \\ &= 22270 - 3843a = 21886a \\ &= \frac{21886}{3} a = 7.295{,}333 \end{aligned}$$

After getting the values of a and b, substitute the values of a and b obtained into the trend moment equation, namely, Y = a + bX. Then, the trend moment formula is obtained, namely Y = 7,295.33 + 128x. To calculate sales predictions in July, enter an x value of 3 into the equation. Then we get a sales prediction in July of 7679, likewise for the next period.

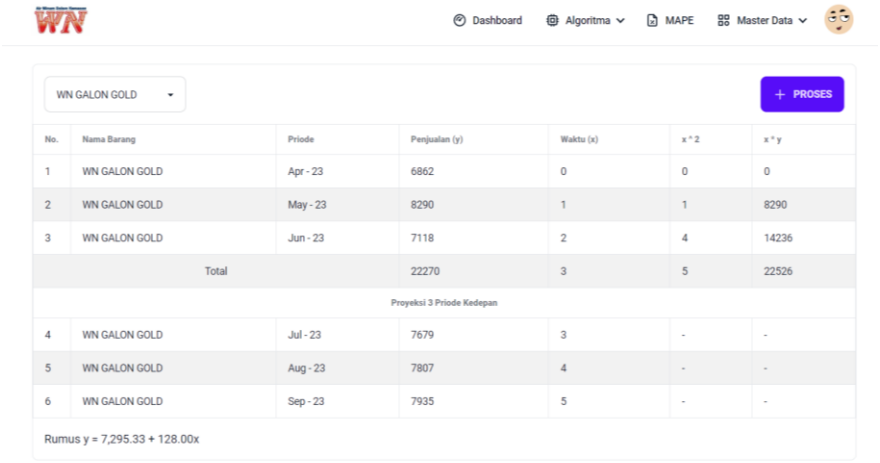


Figure 5. Trend Moment Algorithm Process

3.2. Mean Absolute Percentage Error Result

Mean absolute percentage error is used to test the error rate of forecasting done by the trend moment algorithm. The first step taken is to reduce the actual data value with the forecast value, as in Figure 6. Next, calculate the MAPE value for each period using the formula $\text{MAPE} = \frac{1}{n}\sum\left|\frac{y_t - y'_t}{y_t}\right| \times 100$. The following is the calculation of the error value in July.

$$\text{MAPE} = \frac{1}{n}\sum\left|\frac{y_t - y'_t}{y_t}\right| \times 100 \quad \text{MAPE} = \frac{1}{1}\sum\left|\frac{117}{7796}\right| \times 100 \quad \text{MAPE} = 0{,}015007 \times 100 \quad \text{MAPE} = 1{,}5007$$

I obtained an error value of 1.50% in July. Do the same calculation for the next period. After the error value for each period is obtained. Next, calculate the average error value with the formula $\bar{X} = \frac{\sum MAPE}{n}$.

$$\bar{X} = \frac{\sum MAPE}{n} \qquad \bar{X} = \frac{1{,}50 + 3{,}21 + 0{,}43}{3} \qquad \bar{X} = \frac{5{,}14}{3} \qquad \bar{X} = 1{,}71$$

The MAPE testing results obtained an average error value of 1.71%, so it can be said to be very accurate based on Table 1.

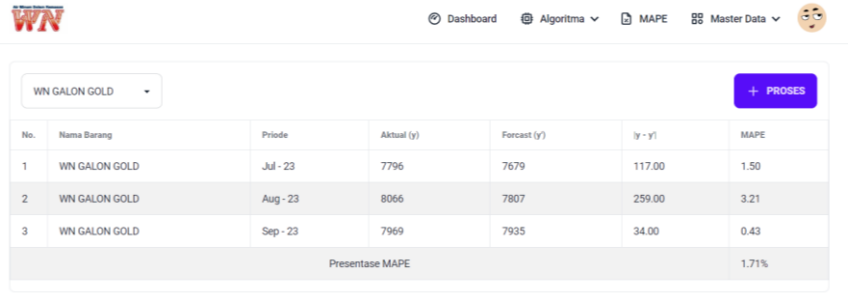


Figure 6. MAPE process

3.3. Black Box Testing Result

Blackbox testing is needed to find out if every function in the system can run properly and reduce errors when the system is running. [15]. The dataset used is in Figure 2. The dataset is entered into a web-based distribution system. The following are the results of system testing using black box testing.

Table 2. Black Box Testing

Feature	Feature Function	Result
Login	Login as admin to access the website	Success
Input, edit, and delete customer or outlet data	Can input, edit, and delete data in the customer or outlet data section	Success
Input, edit, and delete product data	Can input, edit, and delete data in the product data section	Success
Input, edit, and delete sales data	Can input, edit, and delete data in the sales data section	Success
Import xls files in the sales data section	Can import xls files in the sales data section	Success
DRP Process	The system will calculate DRP to customers or outlets based on the inputted data.	Success
Trend moment process	The system will predict the total product sales in the following months.	Success
MAPE Process	The system will calculate the error rate of actual sales data and forecasting results data.	Success

3.4. Discussion

The results of this research indicate that the distribution system website created can run well. The system creation process begins with creating a flowchart, collecting sales data, determining an algorithm that can manage the supply chain, and determining an algorithm that can predict data that shows trends. The algorithms chosen are the distribution requirements planning algorithm and the trend moment algorithm which are applied to a web-based system. Implementation of a web-based distribution system using the DRP algorithm and Trend Moment method brings several benefits, including increased operational efficiency, reduced costs, and improved customer service. However, challenges such as data security, system integration, and telecommunications infrastructure availability still need to be considered. By paying attention to these challenges and taking appropriate steps, companies can optimize their distribution systems to achieve competitive advantage in an increasingly competitive business environment. This research has several limitations that need to be considered so that they can

be improved in further research, namely limited scope, limited data, limited generalization, limited affordability, and limited time. By considering these limitations, it is hoped that future research can evaluate additional factors that may influence the implementation of distribution systems in the future.

4. Conclusion

Based on the results of the distribution system testing using distribution requirement planning algorithms and trend moment algorithms, it can be concluded that the system is functioning well. The system can input, edit, and delete data in customer or outlet data, product data, and sales data. Additionally, the system can carry out DRP calculations as a whole or per customer or outlet. Moreover, the system is capable of forecasting using a trend moment algorithm, which produces an error value of 1.71%, indicating high accuracy. Implementing a web-based distribution system can provide various benefits for companies like PT Anugrah Wina Sentosa. By utilizing the DRP algorithm and the Trend Moment method, companies can improve operational efficiency and respond to market changes. Through this system, companies can manage stock more efficiently, optimize product distribution strategies, and enhance customer service, as evidenced by the results observed in MAPE and black box testing. However, it is essential to address challenges such as data security and system integration to ensure the smooth operation of the distribution system. By addressing these challenges and taking appropriate steps, companies can optimize their distribution systems to achieve a competitive advantage in an increasingly competitive business environment. Nevertheless, several limitations of the study need to be noted. Firstly, there is a limited scope, as this research may not cover all relevant aspects of implementing a web-based distribution system, such as organizational, financial, or operational factors. Secondly, there are data limitations, as this research is based on limited data, thereby limiting the depth of analysis or conclusions that can be drawn. Furthermore, limited generalizability implies that the findings from this study may not be directly applicable to all business or industrial contexts. Additionally, there are limitations in affordability, making it challenging for small or medium-sized companies with limited resources to implement the research findings. Lastly, there are time constraints, as this study may not have been able to capture all relevant time factors. By acknowledging these limitations, future research can be directed to evaluate additional factors that may influence the implementation of distribution systems more comprehensively.

5. Acknowledgments

The researcher would like to thank Tadulako University for the extraordinary opportunity to participate in the Merdeka Belajar Kampus Merdeka (MBKM) program, especially in the Independent Study Program. Researchers appreciate the support, guidance, and inspiring learning environment that has been provided by lecturers, staff, and the entire academic community. This experience has been a valuable journey that provides new insights, skills, and opportunities for the researcher's self-development. Thank you for the dedication of Tadulako University to providing an innovative educational platform and giving researchers the space to develop independently.

References

[1] K. Kulsum, Y. Muharni, and M. R. Mulyawan, "Product distribution scheduling with distribution requirement planning method (Case study of bottled drinking water products)," Tech. J. Science and Technol., vol. 16, no. 1, p. 45, 2020, doi: 10.36055/tjst.v16i1.7800.

[2] Y. Yulyanah and S. Y. Kusumastuti, "Tax Avoidance in Manufacturing Companies in the Goods and Consumption Industry Sector of the Food and Beverage Sub-Sector Listed on the Indonesia Stock Exchange for the Period 2013-2017," Media Ekon., vol. 27, no. 1, pp. 17-36, 2019, doi: 10.25105/me.v27i1.5284.

[3] E. A. Putri, R. Fayaqun, and A. S. P, "Analysis of Rice Distribution Scheduling at Perum Bulog Bengkulu Regional Office with Distribution Requirement Planning (DRP) Method," vol. 7, pp. 16660-16671, 2023.

[4] S. L. Karundeng, Thessa Natasya Mandey and J. S. B. Sumarauw, "Analysis of Timber Distribution Channels (Case Study at Cv. Karya Abadi, Manado)," J. EMBA J. Ris. Econ. Management, Business and Account., vol. 6, no. 3, pp. 1748-1757, 2018.
[5] M. B. Muttaqin, S. Martini, and R. Aurachman, "Design and Scheduling of Household Product Distribution Activities Using the Distribution Requirement Planning (DRP) Method at PT XYZ to Align Product Delivery to Retail," J. System Engineering. Ind., vol. 4, no. 01, p. 56, 2017, doi: 10.25124/jrsi.v4i01.210.
[6] Z. Ulhaq Farid and E. Aristriyana, "Product Distribution Scheduling with the Distribution Requirement Planning (Drp) Method at Cv. Kijang Mas 69 in Ciamis," J. Ind. Galuh, vol. 3, no. 02, pp. 93-98, 2023, doi: 10.25157/jig.v3i02.2976.
[7] D. C. T. Agustin and G. M. Sudiartha, "Optimization of Planning and Scheduling Distribution Activities (Empirical Study at Packed Drinking-Water Distribution Company in Bali, Indonesia)," Academia.Edu, no. 4, pp. 611-616, 2021, [Online]. Available: https://www.academia.edu/download/66858644/ZZZA2154611616.pdf
[8] D. N. E. Ardini, "Sales Forecasting System Using Trend Moment Method at Desktop-Based Nabila Furniture Shop in Paguyangan Brebes," J. Inform. Upgris, vol. 5, no. 2, pp. 130-134, 2019, [Online]. Available: http://journal.upgris.ac.id/index.php/JIU/article/download/4346/2895
[9] A. Kusna et al., "Chart Application Flowcharts Service as a Marketing Strategy of Education in Vocational Secondary School 2 Salahuddin Malang, Indonesia," vol. 508, no. 74, pp. 424-427, 2020, doi: 10.2991/assehr.k.201214.271.
[10] E. Listiana and M. Aziz Muslim, "Application of Adaboost for Support Vector Machine Classification to Improve Accuracy in Diagnosing Chronic Kidney Disease, " Proceedings of SNATIF, no. 2015, pp. 875-881, 2017.
[11] I. Ilyas, F. Marisa, and D. Purnomo, "Implementation of Trend Moment Method (Forecasting) for New Students of Widyagama University Malang," JOINTECS (Journal Inf. Technol. Comput. Sci., vol. 3, no. 2, 2018, doi: 10.31328/jointecs.v3i2.785.
[12] H. D. Budiana, S. Lestanti, and S. Nur Budiman, "Web-based Baby Diaper Stock Prediction Application with Trend Moment Method," JATI (Journal of Mhs. Tech. Inform., vol. 6, no. 1, pp. 26-30, 2021, doi: 10.36040/jati.v6i1.4362.
[13] M. Fahrur Rizal and D. Wahyu Widodo, "Forecasting with the Trend Moment Method to Predict the Number of Sales of Healthy Products at CV. Surya Willis," Semin. Nas. Inov. Technol., pp. 29-34, 2021.
[14] M. A. Muzakki, M. A. Sabila, S. Sundari, and B. Wisnuadhi, "Analysis of Prophet Algorithm to Predict Food Prices in Bandung City," Pros. Ind. Res. Work. Natl. Semin., vol. 12, pp. 659-664, 2021.
[15] M. Jindan, H. R. Ngemba, S. Hendra, R. Laila, and S. Syahrullah, "Integrated Web-based Palu City Blood Donor Service Application Model Using ReactJS and ExpressJS," Adv. Sustain. Sci. Eng. Technol., vol. 5, no. 3, p. 02303013, 2023, doi: 10.26877/asset.v5i3.17196.

Advance Sustainable Science, Engineering and Technology (ASSET)
Vol. 6, No.2, April 2024, pp. 02402012-01 ~ 02402012-08
ISSN: 2715-4211 DOI: https://doi.org/10.26877/asset.v6i2.18323

Supplier Selection Modeling and Analysis in the Metal Casting Industry Using Analytical Hierarchy Process

Zahra Nasywari Firdantara, Qurtubi*, Danang Setiawan

Department of Industrial Engineering, Faculty of Industrial Technology, Universitas Islam Indonesia, Jl. Kaliurang KM 14,5 Sleman, Yogyakarta, Indonesia

*qurtubi@uii.ac.id

Abstract. This study presents a supplier selection problem of a raw material using an analytical hierarchy process. In the existing process, the delivery of raw materials experienced delays and impacted the production process. Therefore, this paper aims to determine the criteria for supplier selection and provide recommendations for the best vendor to be selected from the last request for quotation document. Analytical Hierarchy Process (AHP), as a multi-criteria approach, was utilized in this research, starting from determining criteria, the weight of criteria, and the final score for each supplier. Through discussions with the company's expert, procurement department, and users, as well as a review of the previous studies, this research defined three criteria, each consisting of three sub-criteria. The AHP approach was utilized to evaluate and determine the weights for the three criteria, yielding the following results: quality (62%), price (28%), and delivery (10%). The identified criteria, sub-criteria, and respective weights are subsequently utilized in a supplier selection scenario. Three suppliers of mild steel raw materials were evaluated using the weights of the criteria and sub-criteria obtained. Supplier 1 was selected because of having the higher alternative value of 0.602. The use of AHP in supplier selection is often impractical and contains subjectivity. Therefore, further research can be performed by integrating AHP with other methods, such as weighted scoring, to facilitate further the vendor selection process and integration with other methods, such as fuzzy logic, to reduce subjectivity.

Keywords: AHP, Logistics, Procurement, Supply Chain, Mild Steel Plate.

(Received 2024-02-01, Accepted 2024-03-01, Available Online by 2024-03-08)

1. Introduction

The competition among manufacturing and service industries is increasing rapidly, supported by the development of the science and technology [1]. Companies must provide satisfactory service to consumers and maintain product quality to survive the fierce competition. In addition, to maintain the quality of its product, the company needs to pay attention to the raw materials used in the production process.

The supplier of the raw materials determines the quality of the raw materials. The supplier is a business partner that plays a crucial role in ensuring the availability and quality of raw materials.

Selecting the right supplier according to the company's criteria can generate profit by minimizing costs incurred [2]. Therefore, supplier selection must be performed correctly and carefully to avoid selecting unperformed suppliers [3].

This study used a metal casting company with a make-to-order strategy as a case study. Supplier selection is a critical aspect of the metal casting industry due to its impact on product quality, cost, and productivity [4]. The studied company produces several products, and one of the primary raw materials used is mild steel. In the existing supplier selection method, the decision was solely based on the price of raw material. The supplier selected based on the low-price criteria experienced a delay in delivery, which affected the production process.

Supplier selection can be classified as a multi-criteria decision-making (MCDM) problem, where decision-makers must consider and assess qualitative and quantitative factors [5]. In supplier selection problems, more than one aspect must sometimes be considered [6]. Akarte [7] stated that decision-makers should consider tangible and intangible criteria in supplier selection, such as product development, manufacturing, quality, and cost.

Various methods have been used to address supplier selection problems in the metal industry, such as SWARA and TOPSIS [8], Fuzzy AHP [9], and AHP [10]. In this study, AHP was employed to evaluate the supplier selection problem. A range of studies have demonstrated the effectiveness of the AHP in supplier selection within the metal industry. Tahriri [10] found that AHP can optimize order quantities and reduce the time to select a supplier. Valim [11] and Chhabile [12] compared AHP with other methods, with both studies concluding that AHP is a consistent and effective approach to supplier selection.

Based on the above-mentioned description, this research aims to determine the criteria and sub-criteria in supplier selection. Interview methods and previous literature studies with similar objects were used to determine the criteria and sub-criteria. After the criteria were obtained, the AHP method was used to determine the weight of each criterion and sub-criteria. The weight of criteria contains important information and determines the supplier selection decision. Finally, the criteria and criteria weights were then implemented to evaluate vendor quotations to determine which vendor should be selected by the company.

2. Methods

A range of studies have applied the AHP to the supplier selection problem, emphasizing its ability to consider both qualitative and quantitative criteria. Garoma [13] highlights the importance of AHP in reducing subjectivity and selecting the best vendor. AHP breaks down complex and unstructured problems into components arranged hierarchically. Subjective values are assigned to the components, and the variables that positively influence the situation's outcome are determined [14].

Figure 1 depicts the research flow, consisting of three main steps: the preliminary, data collection and processing, and discussion and conclusion. In the preliminary step, problems and solutions were defined. The author identifies the problems in the company, then addresses them and formulates them. After finding a problem, a literature study was performed by looking for references from books and journals and making observations. The problem raised was about what criteria, sub-criteria, weight of criteria and sub-criteria, and the company's best supplier. The problem was described into its elements, consisting of criteria and alternatives, and arranged into a hierarchical structure.

The data collection and processing stage consists of determining criteria, sub-criteria, and the weight of the criteria. This research used company experts consisting of procurement and users to determine criteria weights through pairwise comparisons. Before the AHP results are used, the consistency ratio

(CR) value is calculated to determine whether it met the standard of less than or equal to 0.1. The pairwise comparison is repeated if the CR calculation results obtained a value of more than 0.1.

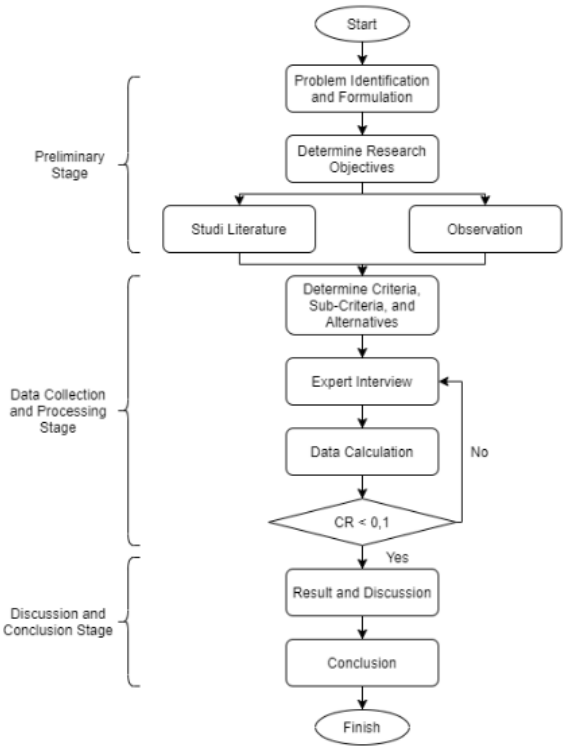


Figure 1. *Research Flow*

3. Results and Discussion

3.1 Determining the Criteria and Sub-Criteria

Many previous studies have applied the AHP approach to select metal casting suppliers and identify the main criteria and sub-criteria. Soepardi [16] and Wang [17] highlight the importance of product quality, cost, and manufacturing capability. In addition, Tahriri [10] and Lin [18] further proposed the trust, net price, and delivery rate as the main criteria. Therefore, previous studies highlight the main criteria: quality, cost, and delivery rate. To ensure relevance to the observed company, this study determined criteria by discussion with related stakeholders such as the procurement department and users. This discussion verifies the importance of the three criteria: quality, cost, and delivery rate. In addition, the discussion also highlights the sub-criteria for each defined criterion.

3.1.1 Quality (A)

Quality refers to a product or service characteristic that supports its ability to satisfy customer needs [19]. Quality represents a company's effort to differentiate its products from competitors to make them more desirable or unique. Quality is used as a criterion because in selecting suppliers, the quality of raw materials affects the quality of the product. Furthermore, the quality criteria are broken down into three sub-criteria as follow:

1. **Quality standard (A1)**
   Quality standards are used to assess and measure the level of quality or adequacy of a product. Quality standards set specific standards or limits that must be met for the product to be considered to meet the desired level of quality.
2. **Product specification conformance (A2)**

The conformity of product specifications refers to the extent to which the product meets the specifications or requirements previously set.

3. **Completeness of documents (A3)**
   Completeness refers to how documents related to checking or testing a product have been prepared and meet predetermined requirements.

3.1.2 Price (B)

A price is a monetary unit exchanged to obtain ownership rights [20]. Price is used as a criterion because competitive prices from suppliers can help companies control production costs to increase companies' profitability. In the price criteria, 3 sub-criteria were defined:

1. **Price fluctuation rate (B1)**
   The level of price fluctuation refers to how often there is a change or instability in the supplier's price. A high level of price fluctuation can cause uncertainty in budget planning and supply chain management.
2. **Price Level (B2)**
   Price level is the price charged by a supplier for a product. It includes the direct price of the purchased goods and additional costs such as shipping or administrative costs.
3. **Price Flexibility (B3)**
   Price flexibility means suppliers are willing to negotiate or provide price adjustments in contractual agreements. Suppliers' flexibility on price can make reaching a mutually beneficial agreement easier.

3.1.3 Delivery (C)

Delivery is part of logistics operations to distribute goods and services from producers to consumers efficiently and accurately [21]. The reason for using delivery as a criterion is that it helps companies mitigate risk, improve operational efficiency, and provide better customer service. In the delivery criteria, there are 3 sub-criteria defined as follows:

1. **Delivery promptness (C1)**
   Delivery promptness refers to the supplier's ability to deliver products according to the agreed schedule.
2. **Lead time (C2)**
   Lead time can be described as the time required from the time of order until the customer receives the product. Suppose the goods are delivered within the delivery tolerance period. In that case, no penalty fee is charged to the vendor, thus providing flexibility for vendors and buyers to overcome the uncertainty of delivery mechanisms and transportation times [22].
3. **Delivery flexibility (C3)**
   Delivery flexibility is to adjust the delivery schedule according to the needs of the customer company. Flexible suppliers can respond quickly to changes in demand or emergencies.

3.2 AHP Structure

Based on the criteria and sub criteria defined in the previous step, Figure 3 illustrated the AHP structured consisting of criteria, sub criteria and alternatives. In the hierarchical structure, the main goal is placed at the first level, the criteria obtained are placed at the second level, the sub-criteria at the third level, and the alternatives at the last level. In this study, the three supplier was used as supplier alternatives and evaluated using pre-defined criteria and sub criteria to get the best supplier.

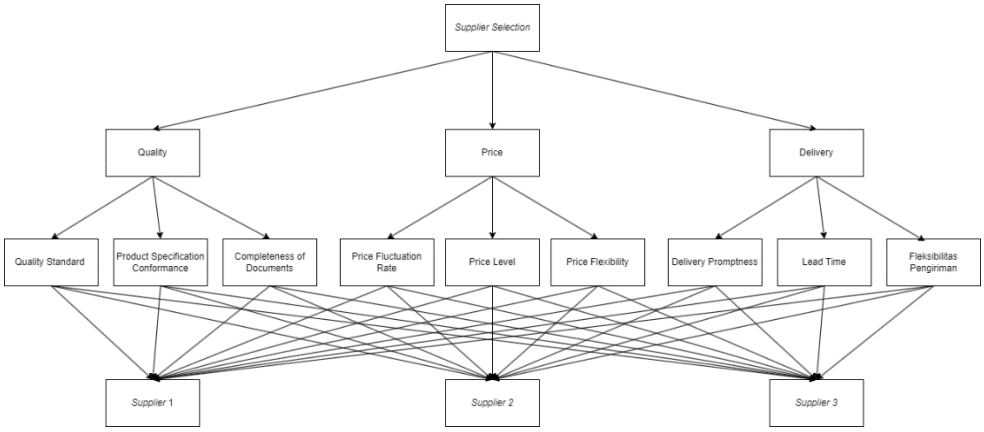


Figure 2. *Hierarchical Structure of Criteria, Sub-criteria and Alternative*

The hierarchical model is subsequently utilized to develop a pairwise comparison questionnaire. This study involved procurement and users as respondents, utilizing the scale presented in Table 1. The consensus approach is used to determine the value of pairwise comparisons.

The Consistency Ratio value of the comparison results was initially assessed using Equation 2. If the CR value is less than or equal to 1, the AHP evaluation can be utilized for subsequent processing. Table 2 illustrates an example of computations for pairwise comparison of criteria. Table 3 provides a summary of the CR values for all pairwise comparisons.

Table 1. *Pairwise Comparison Among Suppliers*

Criteria	Quality	Price	Delivery	Quality	Price	Delivery	1	2	3	4 = 3/2	5 = Sum 4/Sum 1	6 = (5 - Sum 1)/(Sum 1- 1)	7	8 = 6/7
							Total Weight Matrix	***eugen vector***	**Perkalian Matriks**	***Eugen Value***	***λ maks***	***CI***	***IR***	***CR***
Quality	1	3	5	0,6522	0,7059	0,5000	1,8581	0,619352	1,954135	3,155127	3,086690971	0,043345485	0,58	0,074734
Price	1/3	1	4	0,2174	0,2353	0,4000	0,8527	0,284228	0,876357	3,083283				
Delivery	1/5	1/4	1	0,1304	0,0588	0,1000	0,2893	0,096419	0,291347	3,021662				
Total	1,53333333	4,25	10	1	1	1	3	1	3,121839	9,260073				

Table 2 shows the weights for the Quality, Price, and Delivery criteria, which are 0.62, 0.28, and 0.10, respectively. The pairwise comparison results are consistent because the CR is less than or equal to 0.1. Thus, the weighting results can be used for further analysis. The CR evaluation for all pairwise comparisons is shown in Table 3, which indicates that all pairwise comparisons have been consistent.

Table 2. *Consistency Ratio (CR) for all pairwise comparison*

No	Pairwise Comparison	CR
1	Main criteria	0,074
2	Sub criteria: Quality	0,093
3	Sub criteria: Price	0,046
3	Sub criteria: Delivery	0,015
4	Alternative against Sub criteria: Quality	0,081
5	Alternative against Sub criteria: Price	0,015
6	Alternative against Sub criteria: Delivery	0,074

3.3 Result and Discussion

Table 4 presents the result of the AHP approach. Based on the three criteria defined in the previous step, Quality (A) was the criterion with the highest weight (0.62). This number means that the quality aspect is far superior to the other criteria: Price (B) and Delivery (C). In a detailed analysis of Quality criteria, Product specification conformance (A2) is the sub-criteria with the highest weight. This number

indicated that conformity with the predefined specification is the main criterion in supplier selection. The value of Alternative Weight Evaluation Supplier 1 is 0.602, Supplier 2 is 0.288, and Supplier 3 is 0.110. Therefore, it can be concluded that based on these three alternatives, Supplier 1 was chosen because it has the highest Alternative Weight Evaluation value of 0.602.

Table 1 AHP Result

	Attribute									***Alt. Weight Evaluation***
	A			B			C			
	0,62			0,28			0,10			
Attribute weight	*A1*	*A2*	A3	*B1*	*B2*	*B3*	*C1*	*C2*	*C3*	
	0,36	*0,52*	0,12	0,25	0,58	0,16	0,62	*0,24*	*0,13*	
	Alternative									
Supplier 1	0,62	0,61	0,54	0,61	0,62	0,52	0,52	0,60	0,51	0,60
Supplier 2	0,28	0,28	0,34	0,28	0,23	0,33	0,33	0,27	0,36	0,28
Supplier 3	0,09	0,09	0,11	0,09	0,13	0,14	0,14	0,11	0,12	0,11

Further analysis can be performed by evaluating the sensitivity and integrating AHP with another method. Ishak [23] applied AHP in real-world scenarios, using AHP-TOPSIS to determine the best wire rod supplier. AHP can also be integrated with the weighted scoring method in determining the best supplier. AHP is used to determine criteria and criteria weights, while each vendor is given a value for each criterion. The weighted score is then obtained by multiplying the weight of the criteria by the vendor's score. Through this integration, supplier selection becomes easier for companies to do.

4. Conclusion

This research aims to model and analyze supplier selection for metal-casting raw materials using the AHP method. Based on previous research and discussions with procurement and users, the criteria identified are quality, price, and delivery. Each criterion is further elaborated in the three sub-criteria: quality standards, conformity to product specifications, completeness of checking documents, level of price fluctuations, price level, price flexibility, delivery accuracy, lead time, and delivery flexibility. The evaluation results with AHP obtained weights for each main criterion, which are 0.62, 0.26, and 0.10, and are used for supplier selection. The results of supplier selection using AHP showed that supplier 1 is the best supplier with a weighted value of 0.602. The results of the criteria and criteria weights from this research can be used by companies that use metal as raw materials to determine the best supplier. The use of AHP in supplier selection is often impractical and contains subjectivity. Therefore, further research can be performed by integrating AHP with other methods, such as weighted scoring, to further facilitate the vendor selection process and integration with other methods, such as fuzzy logic, to reduce subjectivity.

References

[1] M. Meri, H. Fandeli, R. Linda, And A. M. Rusdi, "Analisis Pemilihan Supplier Oli Menggunakan Metode Analytical Hierarchy Process Di Bengkel Jeffry Motor," Vol. 4307, No. June, Pp. 355–361, 2023.

[2] H. Taherdoost And A. Brard, "Analyzing The Process Of Supplier Selection Criteria And Methods," *Procedia Manuf.*, Vol. 32, Pp. 1024–1034, 2019, Doi: 10.1016/J.Promfg.2019.02.317.

[3] H. K. Simanjorang, Y. Widharto, C. Index, And B. Baku, "Analisis Pemilihan Supplier Bahan Baku Kelapa Sawit Menggunakan Analytical Hierarchy Process (Ahp) Pada," 2008.

[4] M. E. González, G. Quesada, And C. A. M. Monge, "Determining The Importance Of The Supplier Selection Process In Manufacturing: A Case Study," *Int. J. Phys. Distrib. Logist. Manag.*, Vol. 34, No. 6, Pp. 492–504, 2004, Doi: 10.1108/09600030410548550.

[5] Ž. Stević, D. Pamučar, A. Puška, And P. Chatterjee, "Sustainable Supplier Selection In Health Care Industries Using A New Mcdm Method:Measurement Of Alternatives And Ranking According To Compromise Solution (Marcos)," *Comput. Ind. Eng.*, Vol. 140, No. December 2019, P. 106231, 2020, Doi: 10.1016/J.Cie.2019.106231.

[6] L. Johana Hernández Viveros, J. Catalina Murcia, And D. Alfonso López Sarmiento, "Design Of A Multicriteria Model With A Fuzzy Analytical Hierarchy Process-Fahp, For The Selection Of Suppliers," *Int. J. Eng. Res. Technol.*, Vol. 12, No. 12, Pp. 2749–2754, 2019.

[7] M. Akarte, N. Surendra, B. Ravi, And N. Rangaraj, "Web Based Casting Supplier Evaluation Using Analytical Hierarchy Process," *J. Oper. Res. Soc.*, Vol. 52, No. 5, Pp. 511–522, 2001, Doi: 10.1038/Sj.Jors.2601124.

[8] A. K. Narayanan And N. Jinesh, "Application Of Swara And Topsis Methods For Supplier Selection In A Casting Unit," *Int. J. Eng. …*, Vol. 7, No. 05, Pp. 456–458, 2018, [Online]. Available: Https://Www.Academia.Edu/Download/59389323/Application-Of-Swara-And-Topsis-Methods-For-Supplier-Selection-In-A-Casting-Unit-Ijertv7is05021020190525-68108-Vqg7o5.Pdf.

[9] F. Tahriri, "Supplier Assessment And Selection Using Fuzzy Analytic Hierarchy Process In A Steel Manufacturing Company," *J. Sci. Res. Reports*, Vol. 3, No. 10, Pp. 1319–1338, 2014, Doi: 10.9734/Jsrr/2014/8627.

[10] F. Tahriri, M. R. Osman, A. Ali, R. M. Yusuff, And A. Esfandiary, "Ahp Approach For Supplier Evaluation And Selection In A Steel Manufacturing Company," *J. Ind. Eng. Manag.*, Vol. 1, No. 2, Pp. 54–76, 2008, Doi: 10.3926/Jiem.2008.V1n2.P54-76.

[11] L. Valim De Freitas, A. Paula Barbosa Rodrigues De Freitas, E. Vizconde Veraszto, F. Augusto Silva Marins, And M. Borges Silva, "Decision-Making With Multiple Criteria Using Ahp And Maut: An Industrial Application," *Eur. Int. J. Sci. Technol.*, Vol. 2, No. 9, Pp. 93–100, 2013, [Online]. Available: Www.Eijst.Org.Uk.

[12] S. B. Chhabile And R. S. Dalu, "Supplier Selection Using Ahp And F-Ahp Approach : A Case Study," *Int. J. Sci. Res.*, Vol. 3, No. 5, Pp. 1910–1914, 2014.

[13] T. Garoma And S. Diriba, "Modeling And Analysis Of Supplier Selection Method Using Analytical Hierarchy Process (Ahp)," *Sci. Technol. Arts Res. J.*, Vol. 3, No. 1, P. 145, 2014, Doi: 10.4314/Star.V3i1.24.

[14] S. N. Hakim And A. J. Putra, "The Best Location Selection Using Analytical Hierarchy Process Method," *Int. J. Ind. Optim.*, Vol. 3, No. 1, Pp. 68–79, 2022, Doi: 10.12928/Ijio.V3i1.4438.

[15] B. Zhang, W. Pedrycz, A. R. Fayek, And Y. Dong, "A Differential Evolution-Based Consistency Improvement Method In Ahp With An Optimal Allocation Of Information Granularity," *Ieee Trans. Cybern.*, Vol. 52, No. 7, Pp. 6733–6744, 2022, Doi: 10.1109/Tcyb.2020.3035909.

[16] A. Soepardi, . P., M. Chaeron, And I. Anggraini, "Penentuan Kriteria Pemilihan Strategi Sistem Manufaktur Menggunakan Analytic Hierarchy Process," *J. Tek. Ind.*, Vol. 14, No. 2, Pp. 107–114, 2011, Doi: 10.9744/Jti.14.2.107-114.

[17] C. N. Wang, D. C. Dang, Q. Q. Vu, And Y. X. Zeng, "Supplier Selection For Manufacturing Industries," *Proc. 2018 Ieee Int. Conf. Adv. Manuf. Icam 2018*, Pp. 268–271, 2018, Doi: 10.1109/Amcon.2018.8614833.

[18] G. H. Lin, C. A. Chuang, C. L. Tan, S. F. Yeo, And F. Y. Wu, "Supplier Selection Criteria Using Analytical Hierarchy Process (Ahp)-Based Approach: A Study In Refractory Materials Manufacturers," *Ind. Manag. Data Syst.*, Vol. 123, No. 6, Pp. 1814–1839, 2023, Doi: 10.1108/Imds-06-2022-0370.

[19] Irdhayanti, "Pembelian Produk Busana Muslim Pada Toko Rumah Jahit Akhwat (Rja)," *J. Econ. Manag. Account.*, Vol. 1, No. 1, Pp. 49–59, 2020.

[20] D. Suparman, "Pengaruh Lokasi, Harga Dan Promosi Terhadap Keputusan Pembelian Konsumen Di Minimarket Cinagen Kab. Sukabumi," *J. Ekon.*, Vol. 07, No. 1, Pp. 1–15, 2018.

[21] V. Islami And R. Oktaviani, "Analisis Sistem Pengadaan Barang Dalam Rangka Sistem Pengendalian Intern (Studi Kasus : Pt Lintas Nusantara Perdana)," *J. Adm. Bisnis*, Vol. 2, No. 1,

Pp. 10–18, 2022, Doi: 10.31294/Jab.V2i1.1148.
[22] M. Das Roy And B. R. Sarker, "Optimizing A Supply Chain Problem With Nonlinear Penalty Costs For Early And Late Delivery Under Generalized Lead Time Distribution," *Comput. Ind. Eng.*, Vol. 160, No. September 2020, P. 107536, 2021, Doi: 10.1016/J.Cie.2021.107536.
[23] A. Ishak And S. K. Parinduri, "Supplier Selection Using Analytical Hierarchy Process (Ahp) And Technique For Order Preference By Similarity To Ideal Solution (Topsis)," In *1st International Conference On Engineering And Management In Industrial System (Icoemis 2019) Supplier*, 2019, Vol. 173, Pp. 321–328.

Advance Sustainable Science, Engineering and Technology (ASSET)
Vol. 6, No.2, April 2024, pp. 02402013-01 ~ 02402013-09
ISSN: 2715-4211 DOI: https://doi.org/10.26877/asset.v6i2.18466

Prediction of Birth Rate of Babies at Regional Hospitals in Salatiga City for Future Planning Using the Naïve Bayes Algorithm

Adelliya Dewi Saputri[1*], Hendry[1]

[1]Faculty of Technology & Information, Satya Wacana Christian University, Jl. Diponegoro 52-60, Salatiga, Central Java, Indonesia

* adellsz02@gmail.com

Abstract . Birth rates have a significant impact on population growth and large populations can be a burden on development. In the Salatiga City Regional Hospital, the numbers tend to change every year, with the current population density making it a special concern for the City of Salatiga. Therefore, it is hoped that the application of Data Mining Techniques with the Naive Bayes algorithm can help predict the number of births in the future using the RapidMiner Application. In this research, the population used was Population Data from Salatiga City with a total of 989,674 residents. Then the sample used was 4699 babies from the Salatiga City Regional Hospital. All data was taken from 2019 – 2023 by conducting observations, literature studies and documentation. By analyzing the pattern of each variable and testing the training data against the testing data, a calculation was produced which shows the Testing Data Prediction, namely the "High" label with the number 4.77192E-06, with this the predicted result of the Baby Birth Rate in the Salatiga City Regional Hospital which is influenced by Population Density in 2024 it will be even higher.

Keywords : Naïve Bayes, Baby Birth, Population Density, Prediction, Data Mining, Machine Learning

(Received 2024-02-21, Accepted 2024-03-01, Available Online by 2024-03-09)

1. Introduction

Rapid population development refers to the rapid growth of the human population over a certain period of time. One of the things that can influence population development is the birth rate [21] . The birth rate is something that can influence the increase in population growth [16]. The number of births in Indonesia tends to increase every year. Birth rate, also known as birth rate or fertility rate, refers to the number of births that occur in a population within a certain period of time [1] [17].

Salatiga City Regional General Hospital (RSUD) is a government-owned hospital located on Jl. Osamaliki No.19, Mangunsari, Sidomukti District, Salatiga City, Central Java 50721. Optimal population distribution based on a balance of population and power supported by community quality and environmental quality is the government's hope for each region, especially in Salatiga City. The problem is that not all regions experience population balance and the current increase in population needs to be taken into account [14]. This is because the number of babies born at the Salatiga City Regional Hospital is increasing every year [18] .

Data mining is the process of analyzing data from different points of view and forming it into useful

information . In the context of the Salatiga City Hospital, data mining is used to analyze the birth rate of babies [2] . So it can be a reference for overcoming or anticipating population density in the future [15] [20] [22] . . The Naive Bayes algorithm is used by the Salatiga City Regional Hospital to manage population data, especially birth rates. By using data for 2019-2023, including spontaneous births, caesareans, and population, this algorithm carries out a training process to group the data into negative clusters with a value of 0 and positive clusters with a value of 1.

The previous research that has been carried out in predicting the birth rate of babies at regional hospitals using the Naïve Bayes algorithm which is used as a reference in this research, includes:

a. In research conducted by Nur Isnaini Parihah, Sari Hartini, Juarni Siregar (2020) with the title "Prediction of Baby Birth Rates in Tridaya Sakti Village Using the Naïve Bayes Algorithm" regarding the prediction of baby birth rates in Tridaya Sakti Village showed significant results, especially P(X | Description = "High") is 5.63839E-10. The Tridaya Sakti Village Office can use this prediction as a guide for evaluating population data, as well as as a reference for reducing birth rates, such as increasing the socialization of Family Planning which is expected to be implemented to achieve the goal [3].
b. In research conducted by Fajar Romadhon, Adi Suwondo, Hidayatus Sibyan (2022) with the title "Implementation of the Naïve Bayes Algorithm in the Annual Population Growth Prediction Application in Wonosobo Regency". The prediction results show low population growth because the value (P|low) is greater than (P|high). The population is predicted using the average growth value in the training data which has the same prediction results [4].
c. In research conducted by Sinta Ayu Sabilla, Banni Satria Andoko, Moch Zawaruddin Abdullah (2021) with the title "Naïve Bayes Method for Predicting Baby Birth (Case Study: Midwife S Clinic)". The birth age prediction results obtained at Midwife Sulikah can provide an accuracy rate of 94%, precision of 93.33%, and recall of 96.55% from 200 training data and 50 test data using the Naïve Bayes algorithm. This success is proven by consistent test results between manual calculations and applications [5].

Based on the research above, the author discusses a topic that is slightly different from several previous studies. The writer will conduct research about predicting the birth rate of babies at the Salatiga City Regional Hospital using the *Naïve algorithm Bayes Classifier* which is based on research Previously, the *Naive Bayes Classifier algorithm* was suitable for predicting the number of births in the coming year using baby birth data from 2019 – 2023 . After carrying out the training data processing process by grouping the data into negative clusters which have a value of 0 or low and positive values of 1 or high, data is produced whose level of prediction numbers increases from year to year. It is hoped that this research will be useful for the Salatiga City Regional Hospital and also to reduce population density in the coming year.

2. Method

This research was conducted at the Salatiga City Regional General Hospital. To predict an event that is likely to occur in the future based on data from the last 5 years and information from the past, a Naïve *Bayes algorithm can be used* [6]. This algorithm is a machine learning method that uses probability calculations. The basic concept used by Bayes is Bayes' Theorem [7], namely carrying out classification by calculating probability values. Classification is carried out to determine the category of existing data [9]. An advantage of the Naïve Bayes Classifier is that it will only require a small amount of training data to estimate the parameters required for classification [8].

Bayes' theorem is the basic rule of the Naïve Bayes Classifier. The following Bayes' theorem will be presented in the equation.

$$P\ (H \mid X) = \frac{P(X|H)P(H)}{P(X)}$$

[3]

Information:
X = Data sample that has an unknown class (label).
H = Hypothesis that X is class data (label)
P(H) = Probability of hypothesis H
P(X) = Probability of the observed sample data
P(X|H) = Probability of sample data X if it is assumed that the hypothesis is true

Naïve Bayes Formula Numerical Data Type

Gaussian distribution	**Mean**	**Standard Deviation**
$g(x, \mu, \sigma) = \frac{1}{\sqrt{2\pi}.\sigma} exp^{\frac{-(x-\mu)^2}{2\sigma^2}}$	$\mu = \sum_i^n x_i$	$\sigma = \frac{\sqrt{\sum_i^n (x_i - \mu)^2}}{n-1}$
information : g : Normal Distribution x : Variable Value μ : Calculated Average (Mean) σ: Standard Deviation	information : μ : Calculated Average i : Sample Sequence xi : Value of the i-th Sample n : Number of Samples	information : σ : Standard Deviation xi: Value of X to –I μ : Calculated Average n : Number of Samples

[19]

In this study, the population used was from Salatiga City Population data of 989,674 residents [10]. Then the sample used in this research was 4699 babies at the Salatiga City Regional Hospital. Meanwhile, the methodology that will be used in this research is shown in Figure 1 where these steps will be carried out during the research process. As for the detailed methods in this research, there are stages that can be seen as follows:

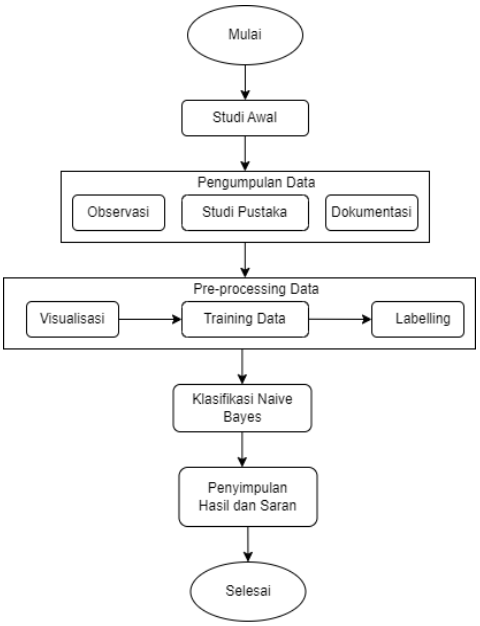


Figure 1. Flowchart of Research Method Stages

This research began with a Preliminary Study at the Salatiga City Regional Hospital to identify birth problems [12]. Data was collected through observation, literature study and documentation. Data processing focuses on identifying problems in RSUD, with visualization using scatter plots for the relationship between population density and baby births. The Kmeans method is used to group data into positive clusters with a value of 1 and negative values with a value of 0 [13]. then the Naïve Bayes algorithm can carry out classification using probability and statistical methods, namely predicting future opportunities based on previous experience. The final stage is Conclusion of Results and Suggestions, where the classification results are used to conclude the birth rate of babies in the Salatiga City Regional Hospital [11]. So that later accurate predictions, suggestions and solutions can be submitted to the Salatiga City Regional Hospital to overcome the potential for increasing population density in the coming year.

3. Results and Discussion

3.1 Data analysis

Naïve Bayes can carry out classification using probability and statistical methods, namely predicting future opportunities based on previous experience. In analyzing Salatiga City Population data using the Naïve Algorithm Bayes begins by determining the training data. After that, the author determines the test data and then uses the Naive Bayes Algorithm to calculate predictions or possibilities.

Table 1. Training Data

Year	Population density	Number of Baby Births	Clusters
2019	195563	1448	Low
2020	196082	1081	Low
2021	196440	662	Tall
2022	200220	683	Tall
2023	201369	825	Tall

After determining the data that will be used as training data, then determine the data that will be used as testing data as shown in the following table.

Table 2. Testing Data

Year	Population density	Number of Baby Births	Clusters
2024	202659	447	?

3.2 Calculations using the Naïve Bayes method

1. Calculating the P(Ci) Value of Class (label) Description

   P(Ci)
   P(Description = Height) 3/5 = 0.6
   P(Description = Low) 2/5 = 0.4

2. Calculating the Mean Value

   Mean

$$\mu = \sum_{i}^{n} x_i$$

   Variable: Population, Description: Height

$$\mu = \frac{196440 + 200220 + 201369}{3} = 199343$$

   Variable: Population, Description: Low

$$\mu = \frac{195563 + 196082}{2} = 195823$$

   Variable: Number of Baby Births, Information: Height

$$\mu = \frac{662 + 683 + 825}{3} = 723{,}3$$

   Variable: Number of Baby Births, Description: Low

$$\mu = \frac{1448 + 1081}{2} = 1114{,}5$$

3. Calculating Standard Deviation Values
   Standard Deviation

$$\boldsymbol{\sigma} = \frac{\sqrt{\sum_i^n (\boldsymbol{x_i} - \boldsymbol{\mu})^2}}{\boldsymbol{n} - \mathbf{1}}$$

Variable: Population, Description: Height

$$\boldsymbol{\sigma} = \frac{\sqrt{(196440 - 119343)^2 + (200220 - 119343)^2 + (201369 - 119343)^2}}{\mathbf{3} - 1} = \mathbf{66347}$$

Variable: Population, Description: Low

$$\boldsymbol{\sigma} = \frac{\sqrt{(\mathbf{195563} - \mathbf{195823})^2 + (\mathbf{196082} - \mathbf{195823})^2}}{\mathbf{2} - \mathbf{1}} = \mathbf{67341}$$

Variable: Number of Baby Births , Information: Height

$$\boldsymbol{\sigma} = \frac{\sqrt{(\mathbf{662} - \mathbf{723,3})^2 + (\mathbf{683} - \mathbf{723,3})^2 + (\mathbf{825} - \mathbf{723,3})^2}}{\mathbf{3} - \mathbf{1}} = \mathbf{6014,14}$$

Variable: Number of Baby Births , Description: Low

$$\boldsymbol{\sigma} = \frac{\sqrt{(\mathbf{1448} - \mathbf{1114,5})^2 + (\mathbf{1081} - \mathbf{1114,5})^2}}{\mathbf{2} - \mathbf{1}} = \mathbf{335,18}$$

1. Calculate the P Value (XK | Ci) for each class i

P(X K | Ci)

$$\boldsymbol{g(x, \mu, \sigma)} = \frac{\mathbf{1}}{\sqrt{\mathbf{2\pi}.\boldsymbol{\sigma}}} \boldsymbol{exp}^{\frac{-(x-\mu)^2}{2\sigma^2}}$$

Variable: Population, Description: Height

$$\boldsymbol{g(x, \mu, \sigma)} = \frac{1}{\sqrt{2\,.3{,}14\,.66347}} exp^{\frac{-(202659-199343)^2}{2.\ 66347^2}} = 0{,}001547274$$

Variable: Population, Description: Low

$$\boldsymbol{g(x, \mu, \sigma)} = \frac{1}{\sqrt{2\,.3{,}14\,.67341}} exp^{\frac{-(202659-195823)^2}{2.\ 67341^2}} = 0{,}001529828$$

Variable: Number of Baby Births , Information: Height

$$g(x, \mu, \sigma) = \frac{1}{\sqrt{2 \times 3{,}14 \times 6014{,}14}} exp^{\frac{-(447-723{,}3)^2}{2 \times 6014{,}14^2}} = 0{,}005140142$$

Variable : Number of Baby Births , Description: Low

$$g(x, \mu, \sigma) = \frac{1}{\sqrt{2 \times 3{,}14 \times 335{,}18}} exp^{\frac{-(447-1114{,}5)^2}{2 \times 335{,}18^2}} = 0{,}00300046$$

2. Calculating the P(X|Ci) Value for each class (label)

   P(X|Ci)

   P (X | Description = “High”)

   =0,001547274 × 0,005140142 = 7.95321E-06

   P (X | Description = “Low”)

   =0,001529828 × 0,00300046 = 4.59019E-06

3. Calculating the Value of P(X| Ci)* P(Ci)

   P (X | Description = “Height”) * P (Description = “Height”)

   = 7.95321E-06 ×0.6 = 4 .77192E-06

   P (X | Description = “Low”) * P (Description = “Low”)
   = 4.59019E-06 × 0.4 = 1.83608E-06

4. Determine the Class Results of the cases that have been calculated the final calculation by multiplying the probability values of this case, the author sees that the value of P ("Low" 1.83608E-06 then the description class (label) is **"High"** .

Table 3. Results of Naïve Bayes calculations

Year	Total Population Density	Number of Baby Births	Information
2024	202659	447	Tall

5. Prediction Results using the *RapidMiner Application RapidMiner* Application calculations show the Testing Data Prediction, namely the information label "High", meaning that the predicted results of the Baby Birth Rate at the Salatiga City Regional General Hospital which is influenced by Population Density in 2024 will be higher.

Table 4. Testing Data Prediction Results from the *RapidMiner Application*

No.	Predictions (Information)	Confidence (Tall)	Confidence (Low)	Year	Number Birth Baby	Amount Density Resident
1	**Tall**	100.00%	0.00%	2024	447	202659

In this study, the prediction level for the birth rate of babies in the Salatiga City Hospital in 2024 was obtained by calculating the probability from the dataset using the Naive Bayes algorithm to produce a description label P(X | Description = "High") = 4.77192E-06. This means that the predicted results of the baby birth rate in the Salatiga City Regional Hospital which is influenced by population density in 2024 will be higher. This is in line with the journal [3] that the results of the prediction of the birth rate of babies in Tridaya Sakti Village in 2020 experienced a high birth rate of 5.63839E-10 using the Naive Bayes Algorithm Method so that by making this prediction, steps can be taken to take action to reduce the birth rate of babies in Tridaya Sakti Village. This is also supported by the journal statement [5] which discusses the prediction of birth age at Midwife Sulikah provides an accuracy rate of more than 85%. Namely, with an accuracy level of 94 %, precision of 93.33% and recall of 96.55% from 200 training data and 50 test data. This test uses test data with the aim of determining the level of precision, recall and accuracy in the Naïve Bayes method.

Conclusion

Using Data Mining can find important information from data warehouses whose benefits have not been known until now. Through the Naïve Bayes algorithm, classification can be carried out using probability and statistical methods, namely predicting future opportunities based on previous experience. The results of the predicted birth rate at the Salatiga City Regional General Hospital in 2024 can be used by the Salatiga City Population Service to assist population officers in evaluating data reports and also as a reference for determining the number of births in the future. The Naive Bayes algorithm can calculate the probability of a dataset. The predicted results for the 2024 birth rate are labeled P(X | Description = "High") = **4.77192E-06** . From the results of this research, the Salatiga City Population Service can take steps or actions to reduce the birth rate by further increasing socialization and family planning development activities.

Acknowledgments

The author would like to thank the Salatiga City Population Service and also the Salatiga City Regional General Hospital for giving permission to the author to carry out observations so that the research runs smoothly. The author also expresses many thanks to Satya Wacana Christian University, especially the Informatics Engineering study program, which has supported this research until completion.

References

[1] Novia, Anggun Elni, "Prediction of Birth Rates in 2022 in North Sumatra Using the Naive Bayes Method," vol. 1, pp. 8-14, 2022.

[2] Idris, Muhammad, "Implementation of Data Mining Using the Naïve Bayes Algorithm to Predict Crime Scenes in Ponorogo Regency," *Paper Knowledge . Toward a Media History of Documents,* vol. 7, no. 1, pp. 1-33, 2020.

[3] Parihah, Nur Isnaini, "Prediction of Infant Birth Rates in Tridaya Sakti Village Using the Naive Bayes Algorithm," *Journal of Students' Research in Computer Science,* vol. 1, no. 2, pp. 77-88, 2020.

[4] Informatics, Engineering, "Implementation of the Naïve Bayes Algorithm in the Wonosobo Application," vol. 3, no. 2, pp. 301-310, 2022.

[5] Sabilla, Sinta Ayu, "Naive Bayes Method for Predicting Baby Birth (Case Study: Midwife S Clinic)," *Polinema Applicative Informatics Seminar,* pp. 2019-2022, 2021.

[6] Syamsiyah, Nur, "Design of a Loan Prediction Information System at the Panca Bhakti Bekasi Cooperative Using the C4.5 Algorithm," *Journal of Science & Technology*, vol. 9, no. 1, pp. 28-43, 2019.

[7] Rizal, Saiful, "Application of the Naïve Bayes Algorithm to Predict New Student Admissions at Al-Amien Vocational School, Wonorejo," *Explore IT: Scientific Journal and Applications of Informatics Engineering*, vol. 10, no. 1, pp. 14-21, 2018.

[8] Qisthiano, M Riski, "Development of a Model for Predicting Student Graduation Rates on Time Using the Naïve Bayes Method," *Budidarma Media Informatics Journal*, vol. 5, no. 3, p. 987, 2021.

[9] Tri Romadloni, Nova, "Comparison of Naive Bayes, Knn and Decision Tree Methods for Krl Commuter Line Transportation Sentiment Analysis," *IKRA-ITH Informatics Journal*, vol. 3, no. 2, pp. 1-9, 2019.

[10] Amin, Nur Fadilah, "General Concepts of Population and Samples in Research," *Pilar Journal*, vol. 14, no. 1, pp. 15-31, 2023.

[11] Indraswari, Naisha Rahma, "Application of Birth Age Prediction Using the Naive Bayes Method," Simetri: Journal of Mechanical Engineering, Electrical and Computer Science, vol. 9, no. 1, pp. 129-138, 2018.

[12] Damar Rani, Handini Arga, "Prediction System for Infant Birth Conditions using Naïve Bayes Classification," *Joined Journal (Journal of Informatics Education)*, vol. 3, no. 2, p. 48, 2020.

[13] Sihombing, Pardomuan Robinson, "Application of Machine Learning Methods in Classifying the Risk of Low Birth Weight Events in Indonesia," MATRIK: Journal of Management, Informatics Engineering and Computer Engineering, vol. 20, no. 2, pp. 417-426, 2021.

[14] Emaniar, Ade Emaniar, "Increasing Maternal Knowledge in Preventing Stunting Through Community-Based Nutrition Education," *Advance Sustainable Science Engineering and Technology*, vol. 5, no. 2, p. 0230205, 2023.

[15] Pusadan, Mohammad Yazdi, "Implementation of Data Mining to Predict the Status of the Labor Process in Pregnant Women Using the Naive Bayes Algorithm," *Technomedia Journal*, vol. 8, no. June 1, pp. 137-153, 2023.

[16] Z. Zulham, "Prediction of Baby Birth Rates in Data Mining Using the Multi Regression Method," *Warta Dharmawangsa*, vol. 17, no. 1, pp. 30-40, 2023.

[17] WC Wahyudin, "Prediction of Stunting in Toddlers in Semarang City Hospitals Using Naive Bayes," vol. 2019, pp. 32-36, 2023.

[18] DFI Lestari, "Factors That Influence the Number of Births in Rapak Dalam Subdistrict," *Innovation*, vol. 14, no. 1, p. 8, 2018.

[19] W. Ananda, "Prediction of Palm Oil Harvest Results Using the Naive Bayes Algorithm," *TIN: Applied Informatics Nusantara Vol*, vol. 1, no. 10, pp. 513-519, 2021.

[20] MT Anwar, "Rain Prediction Using Rule-Based Machine Learning Approach," *Advance Sustainable Science, Engineering and Technology*, vol. 2, no. 1, pp. 1-6, 2020.

[21] VAV Setyawati, "Mapping Exclusive Breastfeeding Coverage And Toddler Stunting Prevalence In Indonesia Based On Web Geographic Information System," *Advance Sustainable Science, Engineering and Technology*, vol. 2, no. 2, pp. 1-5, 2020.

[22] MF Rifai, "Application of the Naïve Bayes Algorithm in the Graduation Level Prediction System for Microsoft Office Specialist (MOS) Certification Participants," *Petir*, vol. 12, no. 2, pp. 131-144, 2019.

Advance Sustainable Science, Engineering and Technology (ASSET)
Vol. 6, No.2, April 2024, pp. 02402014-01 ~ 02402014-08
ISSN: 2715-4211 DOI: https://doi.org/10.26877/asset.v6i2.18530

Development of Website-Based Stunting Prevention Educational Media Services (Case Study: UPT Puskemas Tirto Pekalongan)

Dinita Kusumasari*, Egia Rosi Subhiyakto

Faculty of Computer Science, Universitas Dian Nuswantoro Semarang, Jl. Imam Bonjol No 207 Semarang 50131, Central Java, Indonesia

*111202013038@mhs.dinus.ac.id

Abstract. UPT Puskesmas Tirto is a public health center located in the West Pekalongan sub-district that provides maternal and child health poly services. This poly provides educational services related to the problem of stunting in children through Posyandu activities. However, providing educational services is less effective and efficient because some parents do not have much time to attend counseling so information is not conveyed properly. This research aims to optimize the provision of educational services to parents, prospective mothers, and young women through the use of information technology by developing a website-based stunting prevention educational media service using the ReactJs framework. The development process of this research is carried out by prototyping method. Based on the results of testing using the Blackbox method, it shows that the main features such as nutrition check services, consultation services with Puskesmas nutritionists, and literacy services in the form of articles in the application can run well according to their functions in providing stunting prevention information with 100% results.

Keywords: Website App, Educational Media, Stunting Prevention, Prototyping, React Js

(Received 2024-02-29, Accepted 2024-03-05, Available Online by 2024-03-09)

1. Introduction

Child stunting is one of the health problems that has not been resolved in several countries including Indonesia. In the Central Java region, especially in Pekalongan City, stunting is a serious health problem in children and has not been resolved optimally. This was mentioned in the statement of the Head of Population Control and Family Planning of the Social and Family Planning Agency, who stated that the prevalence rate of child stunting in the Pekalongan City area still reached 23.1% in 2023 [1] considering that the national stunting prevalence standard targeted by 2024 must be below 14% [2] and set by WHO must be below 20% [3].

Based on the results of direct interview observations with one of the maternal and child health clinic officers of the UPT Puskesmas Tirto Pekalongan City, many factors can influence the incidence of stunting in children, this can usually be influenced by several internal factors, including children who

are born from anemic or malnourished mothers, low-calorie intake and nutrition in mothers and children. Then external factors, it is caused by poor parenting patterns, lack of knowledge about basic vaccinations in children, and low levels of knowledge and awareness of parents regarding stunting. Maternal and child health clinic service officers have taken several actions to suppress the internal and external factors that cause stunting in children. One of them is by providing free educational services to parents through counseling related to stunting in children. This educational service is routinely held every month complete with counseling with midwives, nutrition officers, and local Posyandu cadres. However, in reality, there are accessibility constraints faced by parents in attending conventional counseling activities such as busyness and time constraints so that they do not get information about stunting prevention provided during counseling activities. In addition, educational counseling that targets adolescent girls at school is less effective due to limited time and space. This is certainly a serious concern because knowledge about anemia and malnutrition is a basic thing that must be considered in adolescent girls so that later they do not give birth to children with stunting conditions [4]. The findings of these problems can be the basis for the need to optimize stunting prevention education services.

To optimize stunting prevention education services, an update is needed by utilizing technology [5] in the form of educational media services [6] so that information about the chain of stunting problems is conveyed to parents and adolescent girls who will become prospective mothers [7] [8]. The selection of the *website* as a medium for digitizing the delivery of information and education provides a solution because users have the freedom to access it anytime and anywhere through the *browser* on Android and *iOS* devices [9]. Digitalization in this study is defined as the process of transforming health education services from conventional forms to be more interactive and integrated with health workers, to provide convenience and more reliable information for users [10].

Some relevant previous research related to the development of health education service applications includes research [11] that provides KIA (Maternal and Child Health) book education services based on Android applications, with this application making it easier for health cadres to provide health assistance to pregnant women by providing information related to stunting, nutrition-related information, body mass index check calculators, and immunizations. However, some things are lacking in this study, namely the absence of an online counseling system that facilitates the educational assistance process. Furthermore, research [12] developed a service that makes it easier for health workers to monitor the growth data of children affected by stunting. This research only focuses on collecting data on children who are stunted so it does not provide educational features such as article features on stunting prevention. Other research [13] built an android-based health application to facilitate monitoring and prevention of stunting through data inputted by parents. The application developed in this study has been successfully tested on Android version 7 nougat but it is still unknown whether it can be compatible with the latest version of Android. In addition, this research is also not equipped with access to information about stunting and a consultation system so it does not provide educational and interactive content. The development of health education service applications in previous studies has made an important contribution to maternal and child health. However, there are still shortcomings that need to be considered, such as the development of features that specifically focus less on providing educational services on stunting prevention and the absence of interactive features such as online consultations with health workers. In addition, the Android application developed has limitations in accessibility and compatibility with certain specifications.

Seeing the existing problems and shortcomings of previous research, this research was conducted to complement these shortcomings. This research aims to develop a *website-based* stunting prevention educational media service application. This *website* application provides broad accessibility through a digital platform by providing free educational services in the form of access to articles about stunting, body mass index checks on children, and consultation services via *WhatsApp* with Puskesmas nutrition officers so that parents and prospective mothers can access information related to stunting prevention education anytime and anywhere according to their needs, which may not be met by conventional counseling methods. Thus, this web-based approach is expected to provide a more effective solution in overcoming the limitations of stunting prevention education.

2. Research Methods

Application development in this study was developed using the prototyping method. The prototyping method is a process used in software development by involving users directly in the process [14] [15] [16]. User involvement in the software development process aims so that researchers do not misunderstand the system requirements needed by their users [17]. Another goal is that users can monitor each process so that they can understand how the system works that has been created. This is an advantage of the prototyping method compared to other methods such as Rapid Application Development (RAD) or Waterfall [18] [19]. In addition, the use of prototyping methods is very relevant to ongoing projects that require modification because it is responsive to changing user needs [20].

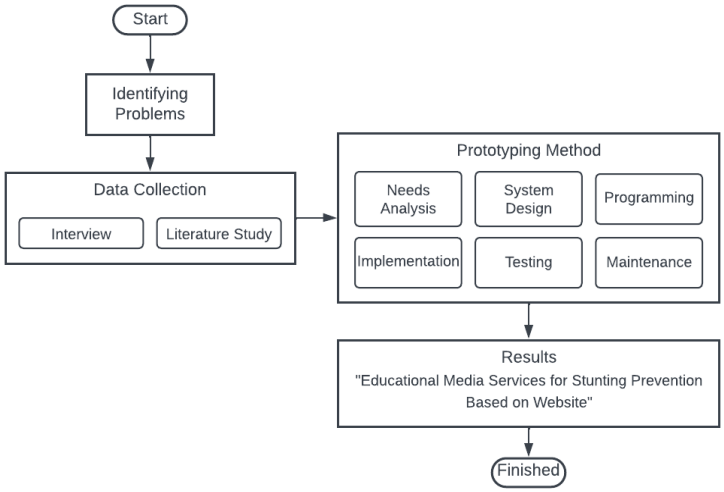


Figure 1. Research stages

Figure 1 illustrates the process of research stages that will be carried out in this study. The stages began with identifying problems related to stunting prevention education services at the mother and child clinic that were not optimal. Research data was obtained through direct interviews with 2 nutrition clinic officers from UPT Puskesmas Tirto, and supported by literature studies such as journals, articles, and relevant sources to understand the need to optimize stunting prevention education services. The questions asked to support the research included factors that cause stunting in children, types of educational services that have been provided to parents, prospective mothers, and adolescent girls, and information related to stunting. The next step is to analyze the interview data in the form of audio recordings and convert them into a table of functional requirements for the features to be developed. In the design stage, a system design will be made which includes use case design and system user interface design. The results of this design are then implemented into code using the JavaScript programming language using the ReactJs framework. After successful implementation, the system will be tested using Blackbox, and maintenance is carried out to ensure the application can run as expected.

3. Results and Discussion

3.1. Planning and Analysis Stage

The results of the analysis of system requirements obtained through direct interview data with relevant health center officers and literature studies, show some of the functional requirements required by the system to meet user needs. The functional requirements of the system to be developed can be seen in Table 1. Furthermore, these functional requirements will be the main guide in developing the application website to ensure that the features produced can provide solutions that meet user needs.

Table 1. Functional requirements

ReqID	Description	Use Case
FPR01	Display the home page	UC1
FPR02	Display the result of the body mass index check	UC2
FPR03	Displaying consultation with officers	UC3
FPR04	Display a selection of stunting education articles	UC4

3.2. System Design Stage

At this stage, the developed website will go through the process of designing a user-friendly interface. This process includes creating user requirements diagrams, and lo-fi, and hi-fi designs according to the system functionality requirements that have been determined from the planning and analysis stages. Where the design process includes determining content such as nutritional information, healthy eating patterns, body mass index checks, and consultation services. The use case diagram in Figure 2 aims to clarify the description of the interaction model between the user and the system, with features such as accessing the home page, checking body mass index, accessing the consultation page after login or registration, and accessing articles related to stunting.

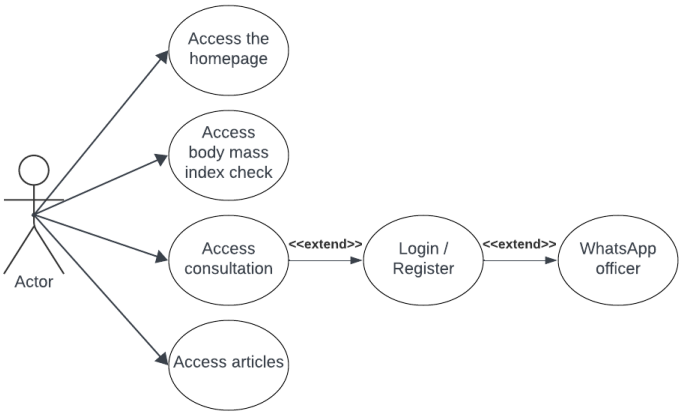


Figure 2. Use case diagram

3.3. Implementation Stage

The following are the results of the website implementation that has been developed. In Figure 3 there is access to the home page which contains an introduction to the SIPENTING website, data on the prevalence of stunting in children according to SSGI, and an overview of stunting articles. In addition, the home page is also equipped with navigation buttons to access the consultation, nutrition check, and article pages.

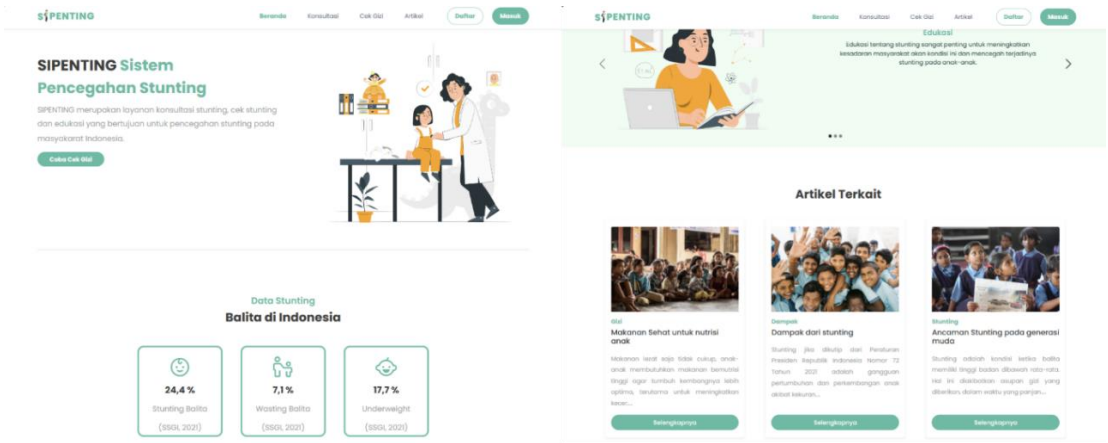


Figure 3. Home page view

Furthermore, Figure 4 displays the results of the implementation of the nutrition check page. On this page, the user can check the nutrition of the child by entering the child's height and weight. The results of the nutrition check are obtained by measuring the body mass index comparison of height and weight of the child.

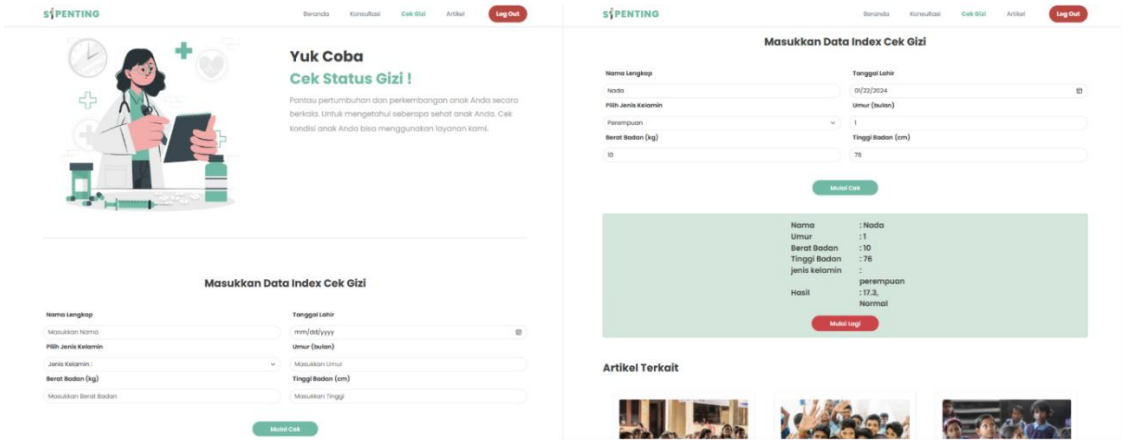


Figure 4. Display of nutrition check page

Figure 5, is a display of the results of the implementation of the consultation page that can be accessed if the user has logged in or registered first as in Figure 6 by entering a username, email, and password. In Figure 6, there is a consultation button which if clicked the user will be directed to the WhatsApp number of the nutritionist of the UPT Tirto Pekalongan City health center who can be contacted if they want to consult about stunting problems in children.

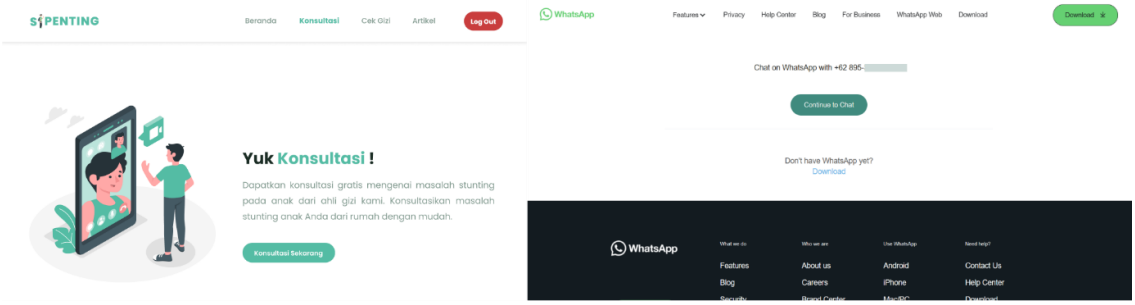


Figure 5. Consultation page view

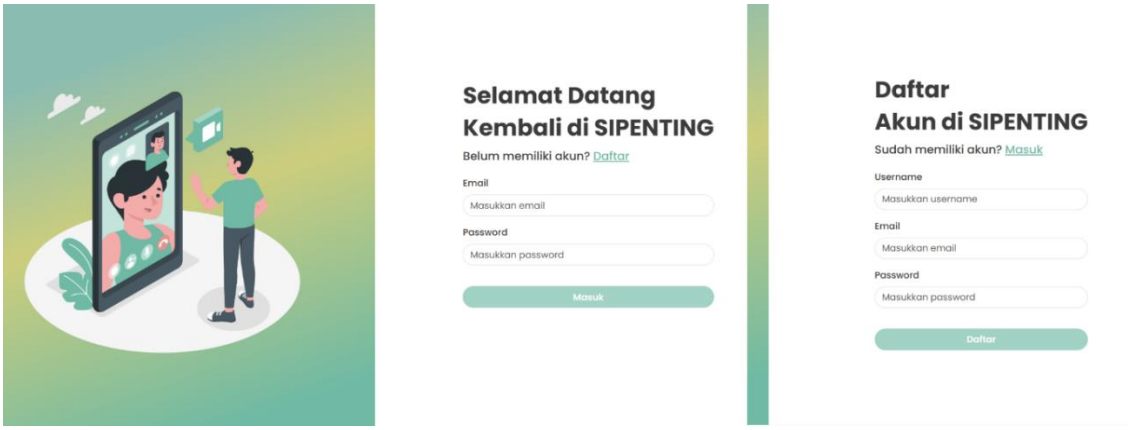


Figure 6. Login and register page view

Figure 7, is a display of the results of the implementation of the article page. On this page, there are several articles provided to improve user literacy and on this page, a button is also provided to filter articles according to the topic the user wants to read. In addition, there is an article detail page to display the content of the article to be read if the user presses the more button.

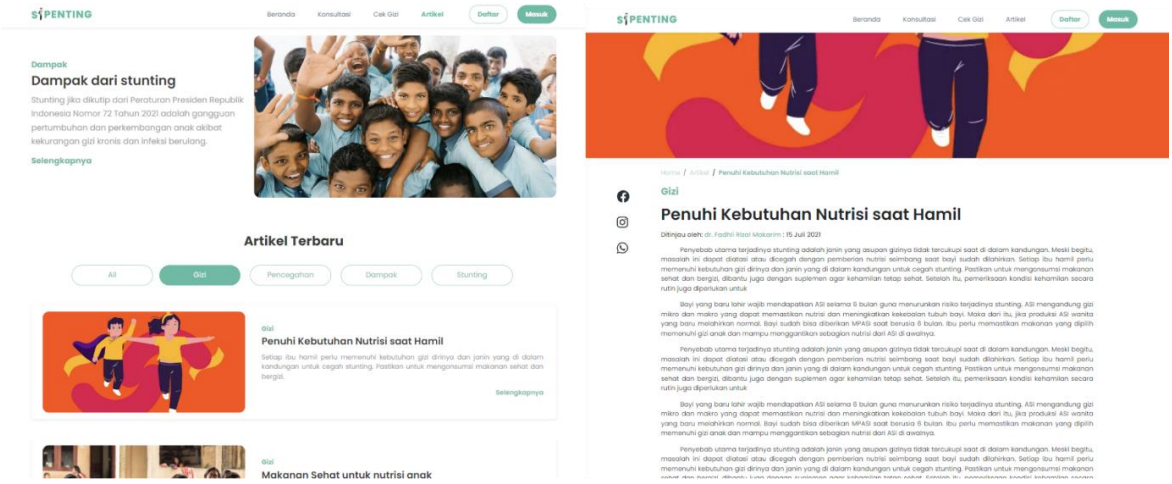


Figure 7. Article page view

3.4. Testing Stage

The successfully implemented SIPENTING *website* requires a test to ensure that the web application runs optimally so that no errors appear on the system. This test is carried out using the Blackbox method by evaluating the input process received and the output given by the system. The test results can be seen in Table 2 below.

Table 2. Blackbox testing results

Features	Test Scenario	Results
Home menu	The user is directed to the home page which contains a brief overview *of* the website, child stunting prevalence data according to SSGI, and an overview of stunting articles.	Successful
Conduct a nutrition check	Users can check nutrition by entering body mass index data such as the child's name, child's gender, child's age child's weight in kilograms, and height in centimeters.	Successful

Conduct a consultation	Users are directed to the login or register page and asked to complete the data. Furthermore, users can consult by clicking the consultation button, and will be forwarded to the WhatsApp of the Puskesmas officer.	Successful
View article	Users can filter for the topic of the article they want to read and can click the More button to be directed to the article detail page to access the content of the article.	Successful

Based on the results of *Blackbox* testing in Table 2, each feature on the SIPENTING *website* successfully fulfills the test scenario and can run optimally as expected. The implications of the findings in this study indicate that limited access to information obtained by parents on stunting prevention can increase the likelihood of stunting in their children. The use of technology can be a foundation for Puskesmas as a public health service to help parents improve their understanding. Utilizing technology such as websites, can encourage further development in the provision of information on stunting prevention, as well as facilitate online consultations with nutrition officers.

4. Conclusion

Based on the results and discussion, it can be concluded that the implementation of stunting prevention educational media services through the SIPENTING website at UPT Puskesmas Tirto has succeeded in improving the effectiveness and efficiency of delivering information. Successfully developed features such as nutrition checks, consultations, and articles make it easier for parents, prospective mothers, and young women to access stunting prevention information flexibly accompanied by nutrition officers through the consultation feature. The existence of this website is expected to help researchers in the future to be able to continue to optimize and update existing features. This research contributes to the field of health education through the use of information technology. Thus, this research not only provides new insights into innovative approaches to delivering health information but also provides a foundation for the development of broader and more equitable stunting prevention strategies in the community.

References

[1] Tim Komunikasi Publik Kota Pekalongan, "Rembug Stunting, Pemkot Komitmen Turunkan Angka Stunting," *Pemerintah Kota Pekalongan*, 2023, [Online]. Available: https://pekalongankota.go.id/berita/rembug-stunting-pemkot-komitmen-turunkan-angka-stunting.html

[2] SSGI, "Hasil Survei Status Gizi Indonesia," *Kementeri. Kesehat. Republik Indones.*, pp. 77–77, 2023, [Online]. Available: https://promkes.kemkes.go.id/materi-hasil-survei-status-gizi-indonesia-ssgi-2022

[3] S. N. Tarmizi, "Prevalensi Stunting di Indonesia Turun ke 21,6% dari 24,4%," *Kementeri. Kesehat. Republik Indones.*, 2023, [Online]. Available: https://sehatnegeriku.kemkes.go.id/baca/rilis-media/20230125/3142280/prevalensi-stunting-di-indonesia-turun-ke-216-dari-244

[4] T. U. Kusuma, "Peran Edukasi Gizi Dalam Pencegahan Anemia Pada Remaja Di Indonesia: Literature Review," *J. Surya Muda*, vol. 4, no. 1, pp. 61–78, 2022, doi: 10.38102/jsm.v4i1.162.

[5] A. A. Fauzi *et al.*, *PEMANFAATAN TEKNOLOGI INFORMASI DI BERBAGAI SEKTOR PADA MASA SOCIETY 5.0.*, no. January. PT. Sonpedia Publishing Indonesia, 2023.

[6] N. Agustian and U. H. Salsabila, "Peran Teknologi Pendidikan dalam Pembelajaran," *ISLAMIKA*, vol. 3, no. 1, pp. 123–133, Jan. 2021, doi: 10.36088/islamika.v3i1.1047.

[7] E. FITRIAMI and A. V. Galaresa, "Edukasi Pencegahan Stunting Berbasis Aplikasi Android Dalam Meningkatkan Pengetahuan Dan Sikap Ibu," *Citra Delima Sci. J. Citra Int. Inst.*, vol. 5, no. 2, pp. 78–85, Dec. 2021, doi: 10.33862/citradelima.v5i2.258.

[8] A. E. Emaniar *et al.*, “INCREASING MATERNAL KNOWLEDGE IN PREVENTING STUNTING THROUGH COMMUNITY-BASED NUTRITION EDUCATION,” *Adv. Sustain. Sci. Eng. Technol.*, vol. 5, no. 2, p. 0230205, Jul. 2023, doi: 10.26877/asset.v5i2.16252.
[9] F. Fattah and H. Azis, “Pemanfaatan Website sebagai Media Penyebaran Informasi pada Desa Tonasa Kecamatan Sanrobone Kabupaten Takalar,” *Ilmu Komput. Untuk Masy.*, vol. 1, no. 1, pp. 15–20, 2020, doi: 10.33096/ilkomas.v1i1.771.
[10] F. Rachmawati *et al.*, *Digitalisasi dalam Perawatan Kesehatan*, vol. 3, no. Juni. WIDINA BHAKTI PERSADA BANDUNG, 2022.
[11] A. P. A. Gita, N. T. Surya, and A. Setyaningsih, “Aplikasi stunting berbasis android guna mempercepat deteksi dini kejadian stunting,” *J. Public Heal. Innov.*, vol. 3, no. 02, pp. 142–150, Jun. 2023, doi: 10.34305/jphi.v3i02.714.
[12] H. A. D. S. Silalahi, M. I. Hutapea, and J. Jamaluddin, “DIGITALISASI AKSES SISTEM LAYANAN TERPADU UNTUK ANAK STUNTING DI DINAS PENGENDALIAN PENDUDUK DAN KELUARGA BERENCANA DI KOTA MEDAN BERBASIS MOBILE,” *TAMIKA J. Tugas Akhir Manaj. Inform. Komputerisasi Akunt.*, vol. 3, no. 1, pp. 17–23, Jun. 2023, doi: 10.46880/tamika.Vol3No1.pp17-23.
[13] A. Hendryani and E. Susana, “Pengembangan Aplikasi Mobile Health Berbasis Android untuk Monitoring dan Evaluasi Stunting,” *J. Sehat Mandiri*, vol. 15, no. 1, pp. 24–32, Jun. 2020, doi: 10.33761/jsm.v15i1.188.
[14] A. Fiorenza and H. Tolle, “Pengembangan Aplikasi Mobile sebagai Media Edukasi Kesehatan Gigi menggunakan Teknologi Firebase serta Metode Prototyping (Studi Kasus RSGM Universitas Brawijaya),” vol. 7, no. 1, pp. 258–266, 2023, [Online]. Available: http://j-ptiik.ub.ac.id
[15] M. Fadhli and Annisa Marion, “Penerapan Metode Prototyping Pada Aplikasi Sentra Pelayanan Kepolisian Terpadu Berbasis Web,” *J. Appl. Comput. Sci. Technol.*, vol. 3, no. 1, pp. 127–133, 2022, doi: 10.52158/jacost.v3i1.267.
[16] S. D. Putra, T. F. Eldiana, and D. Aryani, “Model Pengembangan Aplikasi Mobile E-Dakwah Di Masa Pandemi Covid-19 Dengan Metode Prototyping,” *J. Inf. Syst. Informatics Comput.*, vol. 4, no. 1, pp. 116–121, 2020.
[17] N. Andini, R. Taufiq, D. Y. Priyanggodo, and Y. Sugiyani, “PENGGUNAAN METODE PROTOTYPE PADA PENGEMBANGAN SISTEM INFORMASI IMUNISASI POSYANDU,” *JIKA (Jurnal Inform.*, vol. 7, no. 4, p. 431, Nov. 2023, doi: 10.31000/jika.v7i4.9329.
[18] V. Adi Kurniyanti and D. Murdiani, “Perbandingan Model Waterfall Dengan Prototype Pada Pengembangan System Informasi Berbasis Website,” *J. Syntax Fusion*, vol. 2, no. 08, pp. 669–675, Aug. 2022, doi: 10.54543/fusion.v2i08.210.
[19] T. Pricillia and Zulfachmi, “Perbandingan Metode Pengembangan Perangkat Lunak (Waterfall, Prototype, RAD),” *J. Bangkit Indones.*, vol. 10, no. 1, pp. 6–12, Mar. 2021, doi: 10.52771/bangkitindonesia.v10i1.153.
[20] Y. Wahyudin and D. N. Rahayu, “Analisis Metode Pengembangan Sistem Informasi Berbasis Website: A Literatur Review,” *J. Interkom J. Publ. Ilm. Bid. Teknol. Inf. dan Komun.*, vol. 15, no. 3, pp. 26–40, Oct. 2020, doi: 10.35969/interkom.v15i3.74.

Advance Sustainable Science, Engineering and Technology (ASSET)
Vol. 6, No.2, April 2024, pp. 02402015-01 ~ 02402015-11
ISSN: 2715-4211 DOI: https://doi.org/10.26877/asset.v6i2.18506

Utilizing Data Mining Techniques to Analysis Changes in Purchase Behavior of Batik's Customers

Danang Setiawan[1*], Lisa Alfiyani[2], Joko Sulistio[3], Qurtubi[4]

[1]Industrial Engineering Department, Faculty of Industrial Technology, Universitas Islam Indonesia, Jl. Kaliurang km 14.5, Sleman, Yogyakarta 55584, Indonesia

*danang.setiawan@uii.ac.id

Abstract. Sales transaction data contains rich information potentially used to support company competitiveness. However, interpreting and utilizing transaction data in developing marketing strategies remains a challenge, even for big companies. Therefore, this research aims to develop marketing strategies using data mining techniques. A medium-sized company focusing on producing and selling traditional motif clothes (batik) will be used as a case study. The negative sales trend is the biggest issue currently faced by the company. Hypothetically, this problem is caused by imported products sold at lower prices or changing consumer behavior after pandemic covid. Currently, the company only implements simple analysis of its transaction data. The analysis of transaction data, conducted through five data mining stages, yielded a shift from purchasing small quantities to larger quantities, increased purchases during the final week of each month, and increased purchases on religious occasions. Furthermore, the analysis revealed that 31.29% of all transactions were attributed to loyal consumers, and 192 customers exhibited in Cluster 1 (high transaction quantities and high transaction values). Further investigation also revealed that customers categorized as loyal customers and Cluster 1 have different behaviors that can be used to develop further customer relationship programs. Future research can be conducted by employing data mining techniques to study the organization's assortment of products. Management discussions reveal that changes in consumer buying behavior extend to the selection of items and batik themes.

Keywords: author guidelines, ASSET, article template

(Received 2024-02-27, Accepted 2024-03-09, Available Online by 2024-04-19)

1. Introduction

The retail industry contains abundant transaction data that provide both opportunities and challenges. Transaction data includes essential information that can provide valuable insights into customer preferences and behavior to support the company's competitiveness and customer satisfaction [1]. However, getting useful information from the transaction data is a challenge even for big companies because the previously unstructured data must be processed first [2]. In addition, organizational, cultural, and technological barriers hinder data mining and business intelligence tools in the retail industry [3].

For this reason, up until now, not all industries have successfully managed their transaction data to improve the company's competitiveness or increase customer satisfaction.

Previous research implements different data mining techniques in other cases of utilizing transaction data. Transaction data representing consumer shopping behavior can be used to formulate promotional strategies [1] [4]. Samuel and Gwendolyn [5] studied by integrating Market Basket Analysis (MBA) and 4P strategy (product, price, place, and promotion) in the retail industry. Panjaitan et al. [6] utilized the MBA to develop promotional strategies in product bundling. In this study, data mining techniques will be implemented to analyze transaction data, where the output of analyzing data will then be used to determine the marketing strategy.

This study used a medium-sized company that produces and sells Indonesian traditional motif clothes called *batik* as a case study. The company faces tremendous challenges due to changing consumer purchase behavior and abundant imported products that offer clothes with almost the same motif but far lower prices. Previous research also stated that after the COVID-19 pandemic, the way consumers fulfilled their needs changed significantly, where the clothing industry was the industry that received the most significant effect of the COVID-19 [7]. The pandemic has accelerated the shift towards e-commerce, with customers increasingly relying on online shopping and e-payment methods [8]. Eger et al. [9] stated that retailers and suppliers must successfully consider the new knowledge of consumer behavior and apply it to their selling strategy.

Nevertheless, it is hypothetic that the observed company has loyal customers because the company has been established for over ten years. Therefore, one of the goals of this study is to identify loyal customers by customer segmentation and finding recurring buying behavior. The result of this study, combined with a discussion with the management team, will then be used to determine the marketing strategy.

Several research studies focused on analyzing the behavior of batik consumers in purchasing. Rahadi et al. [10] researched the purchasing behavior of batik customers using a questionnaire. This study found that consumers prefer online and word-of-mouth promotion and select batik designs that combine traditional and modern motifs [10]. Ardiansyah and Febrianti [11] researched customer purchase decisions for batik products using path analysis or regression analysis. The study found that customer engagement influences brand attachment, while brand attachment also influences the purchase decision [11]. The research mentioned above uses direct observation techniques to gain insight into consumer shopping behavior.

In addition to research, research is also conducted using data mining techniques to evaluate consumer shopping patterns. Research by Mulyawan et al. [12] uses clustering techniques using k-means clustering to analyze the motives that consumers are interested in. Clustering techniques are also used by research by Salbinda et al. [13] to group the clothes that customers most demand. Meanwhile, Mardalius and Christy's [14] used clustering techniques to group consumer areas based on sales transaction value.

Based on the abovementioned research, no research comprehensively uses data mining techniques to analyze the sales of batik companies. Therefore, this study aims to identify loyal customers using Pareto, churn analysis, and clustering. The loyal customer identified will be the focus of the company's customer engagement program. This research aims to contribute to the knowledge of the batik business in Indonesia by specifically examining sales transaction data through data mining techniques. The objective is to enhance the company's competitiveness and improve customer satisfaction.

2. Methods

2.1. Data mining technique

Data mining is a technique used to transform unstructured data into valuable knowledge. Data mining can be classified into five techniques: (1) classification, (2) association, (3) prediction, (4) clustering, and (5) outlier analysis [2]. These techniques are advantageous for analyzing large datasets and have been applied across healthcare, finance, and telecommunications. Classification is a method of

classifying data into predefined classes that will become a model of a training set with known class labels. The association is used to identify a frequently used product from data. Prediction is a method to predict how an attribute in data will behave. It is mainly used to identify a pattern within the data. Clustering is a method to group data objects into clusters or groups. Outlier analysis is a phase in which data has a pattern that does not conform to a well-defined notion of normal behavior.

Data mining consists of a series of processes, starting from business understanding, data understanding, data preparation, modeling, evaluation, and deployment [15] [16]. Business understanding is the phase of determining the objectives of the data mining project from a business point of view. Data understanding is the data collection phase and continues with data preparation, consisting of data selection, cleaning, integration, and transformation. After the data is transformed, the modeling phase occurs, where the data is modeled based on the chosen method and parameters. Evaluation is the phase where the researcher evaluates the model results from the modeling phase by validating the model results. Deployment is the phase where the knowledge is obtained through a pattern deployed for the objectives [15]. The following process is selecting a model or analysis technique and evaluating and deploying the results.

2.2. Research flow

This research will be conducted by following the general steps of the data mining technique. Anshu [15] states that data mining is a process that cannot be completed in a single step. Figure 1 illustrates the data mining process used in this study.

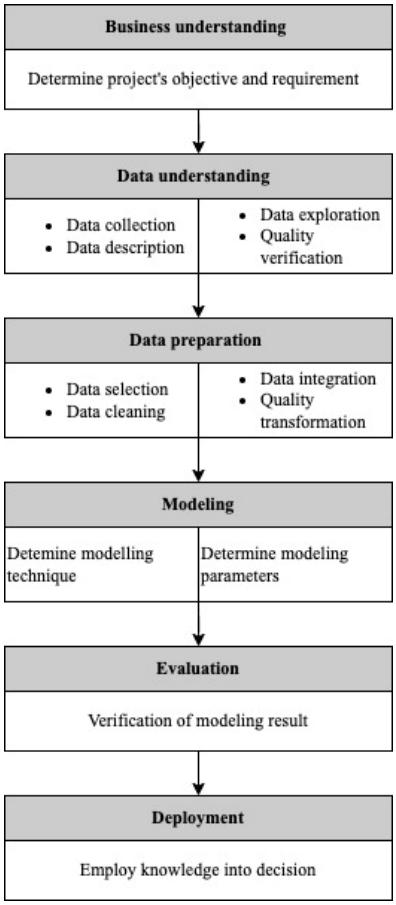


Figure 1. Research step

The description of each step is detailed as follows:

1. Business understanding

Business understanding aims at determining the objectives and requirements of the project. In this study, discussion with the management teams will be conducted to gather the project's goals and needs. Initial meeting obtained that the company's owner wanted to get input for marketing strategies by analyzing current transaction data.

2. Data understanding
   Data understanding consists of several activities: data collection, description, exploration, and quality verification. In this study, data collection will be carried out by downloading sales transaction data through web-based information systems owned by the company. Downloaded data will be verified first with the sales team and the company's owner. Because the pandemic happened from the beginning of 2020 until 2022, the focus group discussion with the management team will be conducted to determine whether data during the COVID-19 pandemic will be used in the next step.
3. Data preparation
   Data preparation contains activities for constructing the final data set to be modeled in the next step. This phase consists of selecting data, cleaning data, data integration, and data transformation.
4. Modeling
   In the modeling phase, the modeling technique and parameters are determined based on the project's objective. The data mining techniques can be classified into classification, clustering, prediction, association rule, neural networks, time series analysis, summarization, and sequence discovery [15].
5. Evaluation
   In the evaluation step, the result of the modeling phase is verified in the context of achieving business objective. The decision in this phase will be to continue to the next stage or review the previous step.
6. Deployment
   In the deployment phase, knowledge obtained is used to generate decisions. In this study, the decision will be the marketing strategy.

3. Results and Discussion

3.1. Data Mining Step

This chapter is presented based on 6 (six) steps of data mining as mentioned in the chapter methodology. However, due to time limitations, this study will cover the 5 (five) steps as follows:

1. Business understanding

Business understanding is a step to determine the objective of a project. In this study, as mentioned in the introduction chapter, the company currently faces tremendous challenges due to changes in customer behavior after conceived-19 pandemic. However, because the company has already been established for more than ten years, it is hypothetic that the observed company has loyal customers. Therefore, this study aims to understand the customers' purchase behavior and identify the company's loyal customers.

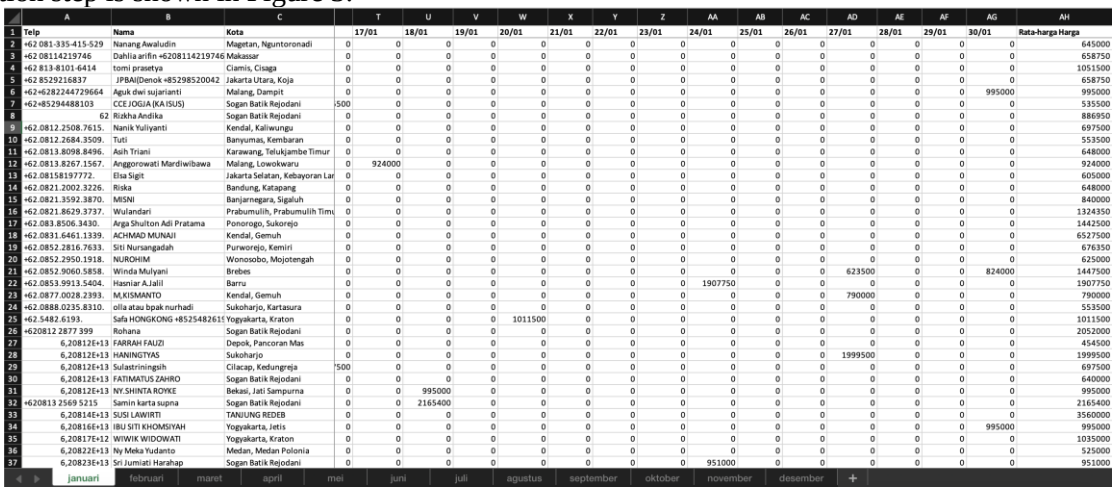
2. Data understanding

Data was collected by downloading the point-of-sales data from the company's information system. The company currently has a web-based platform to manage the transaction data. However, the company has not recorded the customer ID of their customer. Therefore, as seen in Figure 1, there was no standard for inputting the phone number. The data downloaded from the company's information system contain information on each customer's transactions across months in one year. Therefore, this study aims to

understand the purchase pattern and identify the customer with the priority on customer engagement program.

3. Data preparation

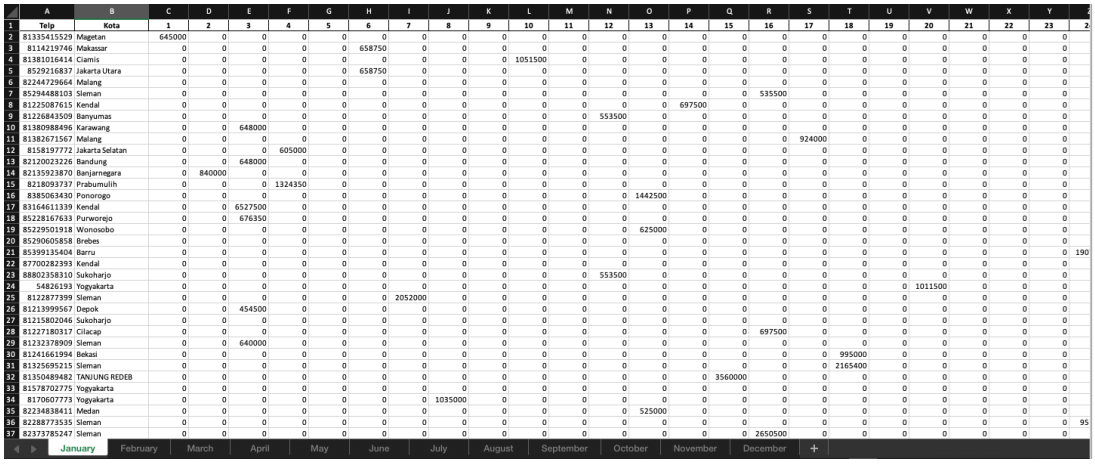
The data downloaded consists of 12 sheets that represent monthly transactions. In total, there were 3900 transactions in 12 months. The initial dataset can be seen in Figure 2, whereas the dataset after the preparation step is shown in Figure 3.



	A	B	C	T	U	V	W	X	Y	Z	AA	AB	AC	AD	AE	AF	AG	AH
1	Telp	Nama	Kota	17/01	18/01	19/01	20/01	21/01	22/01	23/01	24/01	25/01	26/01	27/01	28/01	29/01	30/01	Rata-harga Harga
2	+62 081-335-415-529	Nanang Awaludin	Magetan, Nguntoronadi	0	0	0	0	0	0	0	0	0	0	0	0	0	0	645000
3	+62 08114219746	Dahlia arifin +6208114219746	Makassar	0	0	0	0	0	0	0	0	0	0	0	0	0	0	658750
4	+62 813-8101-6414	tomi prasetya	Ciamis, Cisaga	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1051500
5	+62 8529216837	JPBAI(Denok +85298520042	Jakarta Utara, Koja	0	0	0	0	0	0	0	0	0	0	0	0	0	0	658750
6	+62+6282244729664	Aguk dwi sujarianti	Malang, Dampit	0	0	0	0	0	0	0	0	0	0	0	0	0	995000	995000
7	+62+85294488103	CCE JOGJA (KA ISUS)	Sogan Batik Rejodani	500	0	0	0	0	0	0	0	0	0	0	0	0	0	535500
8	62	Rizkha Andika	Sogan Batik Rejodani	0	0	0	0	0	0	0	0	0	0	0	0	0	0	886950
9	+62.0812.2508.7615.	Nanik Yuliyanti	Kendal, Kaliwungu	0	0	0	0	0	0	0	0	0	0	0	0	0	0	697500
10	+62.0812.2684.3509.	Tuti	Banyumas, Kembaran	0	0	0	0	0	0	0	0	0	0	0	0	0	0	553500
11	+62.0813.8098.8496.	Asih Triani	Karawang, Telukjambe Timur	0	0	0	0	0	0	0	0	0	0	0	0	0	0	648000
12	+62.0813.8267.1567.	Anggorowati Mardiwibawa	Malang, Lowokwaru	0	924000	0	0	0	0	0	0	0	0	0	0	0	0	924000
13	+62.08158197772.	Elsa Sigit	Jakarta Selatan, Kebayoran Lar	0	0	0	0	0	0	0	0	0	0	0	0	0	0	605000
14	+62.0821.2002.3226.	Riska	Bandung, Katapang	0	0	0	0	0	0	0	0	0	0	0	0	0	0	648000
15	+62.0821.3592.3870.	MISNI	Banjarnegara, Sigaluh	0	0	0	0	0	0	0	0	0	0	0	0	0	0	840000
16	+62.0821.8629.3737.	Wulandari	Prabumulih, Prabumulih Timu	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1324350
17	+62.083.8506.3430.	Arga Shulton Adi Pratama	Ponorogo, Sukorejo	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1442500
18	+62.0831.6461.1339.	ACHMAD MUNAJI	Kendal, Gemuh	0	0	0	0	0	0	0	0	0	0	0	0	0	0	6527500
19	+62.0852.2816.7633.	Siti Nursangadah	Purworejo, Kemiri	0	0	0	0	0	0	0	0	0	0	0	0	0	0	676350
20	+62.0852.2950.1918.	NUROHIM	Wonosobo, Mojotengah	0	0	0	0	0	0	0	0	0	0	0	0	0	0	625000
21	+62.0852.9060.5858.	Winda Mulyani	Brebes	0	0	0	0	0	0	0	0	0	0	623500	0	0	824000	1447500
22	+62.0853.9913.5404.	Hasniar A.Jalil	Barru	0	0	0	0	0	0	0	1907750	0	0	0	0	0	0	1907750
23	+62.0877.0028.2393.	M,KISMANTO	Kendal, Gemuh	0	0	0	0	0	0	0	0	0	0	790000	0	0	0	790000
24	+62.0888.0235.8310.	olla atau bpak nurhadi	Sukoharjo, Kartasura	0	0	0	0	0	0	0	0	0	0	0	0	0	0	553500
25	+62.5482.6193.	Safa HONGKONG +8525482615	Yogyakarta, Kraton	0	0	0	1011500	0	0	0	0	0	0	0	0	0	0	1011500
26	+620812 2877 399	Rohana	Sogan Batik Rejodani	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2052000
27	6,20812E+13	FARRAH FAUZI	Depok, Pancoran Mas	0	0	0	0	0	0	0	0	0	0	0	0	0	0	454500
28	6,20812E+13	HANINGTYAS	Sukoharjo	0	0	0	0	0	0	0	0	0	0	1999500	0	0	0	1999500
29	6,20812E+13	Sulastriningsih	Cilacap, Kedungreja	500	0	0	0	0	0	0	0	0	0	0	0	0	0	697500
30	6,20812E+13	FATIMATUS ZAHRO	Sogan Batik Rejodani	0	0	0	0	0	0	0	0	0	0	0	0	0	0	640000
31	6,20812E+13	NY.SHINTA ROYKE	Bekasi, Jati Sampurna	0	0	995000	0	0	0	0	0	0	0	0	0	0	0	995000
32	+620813 2569 5215	Samin karta supna	Sogan Batik Rejodani	0	0	2165400	0	0	0	0	0	0	0	0	0	0	0	2165400
33	6,20814E+13	SUSI LAWIRTI	TANJUNG REDEB	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3560000
34	6,20816E+13	IBU SITI KHOMSIYAH	Yogyakarta, Jetis	0	0	0	0	0	0	0	0	0	0	0	0	0	995000	995000
35	6,20817E+12	WIWIK WIDOWATI	Yogyakarta, Kraton	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1035000
36	6,20822E+13	Ny Meka Yudanto	Medan, Medan Polonia	0	0	0	0	0	0	0	0	0	0	0	0	0	0	525000
37	6,20823E+13	Sri Jumiati Harahap	Sogan Batik Rejodani	0	0	0	0	0	0	0	951000	0	0	0	0	0	0	951000

januari | februari | maret | april | mei | juni | juli | agustus | september | oktober | november | desember

Figure 2. Initial dataset



	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y
1	Telp	Kota	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
2	81335415529	Magetan	645000	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
3	8114219746	Makassar	0	0	0	0	0	658750	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
4	81381016414	Ciamis	0	0	0	0	0	0	0	0	0	1051500	0	0	0	0	0	0	0	0	0	0	0	0	0
5	8529216837	Jakarta Utara	0	0	0	0	0	658750	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
6	82244729664	Malang	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
7	8529448103	Sleman	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	535500	0	0	0	0	0	0	0
8	81225087615	Kendal	0	0	0	0	0	0	0	0	0	0	0	0	0	697500	0	0	0	0	0	0	0	0	0
9	81226843509	Banyumas	0	0	0	0	0	0	0	0	0	0	0	553500	0	0	0	0	0	0	0	0	0	0	0
10	81380988496	Karawang	0	0	648000	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
11	81382671567	Malang	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	924000	0	0	0	0	0	0
12	8158197772	Jakarta Selatan	0	0	0	605000	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
13	82120023226	Bandung	0	0	648000	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
14	82135923870	Banjarnegara	0	840000	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
15	8218093737	Prabumulih	0	0	0	1324350	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
16	8385063430	Ponorogo	0	0	0	0	0	0	0	0	0	0	0	0	1442500	0	0	0	0	0	0	0	0	0	0
17	8316461339	Kendal	0	0	6527500	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
18	85228167633	Purworejo	0	0	676350	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
19	85229501918	Wonosobo	0	0	0	0	0	0	0	0	0	0	0	0	625000	0	0	0	0	0	0	0	0	0	0
20	85290605858	Brebes	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
21	85399135404	Barru	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
22	87700282393	Kendal	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
23	8880235810	Sukoharjo	0	0	0	0	0	0	0	0	0	0	0	553500	0	0	0	0	0	0	0	0	0	0	0
24	54826193	Yogyakarta	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1011500	0	0	0
25	8122877399	Sleman	0	0	0	0	0	0	2052000	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
26	81213999567	Depok	0	0	454500	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
27	81215802046	Sukoharjo	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
28	81227180317	Cilacap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	697500	0	0	0	0	0	0
29	81232378909	Sleman	0	0	640000	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
30	81241661994	Bekasi	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	995000	0	0	0	0	0
31	81325695215	Sleman	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2165400	0	0	0	0	0
32	81350489482	TANJUNG REDEB	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3560000	0	0	0	0	0	0	0	0
33	8157802775	Yogyakarta	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
34	8170607773	Yogyakarta	0	0	0	0	0	0	0	1035000	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
35	82234838411	Medan	0	0	0	0	0	0	0	0	0	0	0	0	525000	0	0	0	0	0	0	0	0	0	0
36	82288773535	Sleman	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
37	82373785247	Sleman	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2650500	0	0	0	0	0	0	0

January | February | March | April | May | June | July | August | September | October | November | December

Figure 3. Dataset after preparation step

The data preparation consists of the following steps:

a. Phone number standardization. The phone number is standardized using the following action:
   - Delete the + (plus), - (minus), and . (dot).
   - Change 62 at the beginning of with 0, where 62 is the country code for Indonesia.
   - Delete leading zero; therefore, all phone numbers start with non-zero values.

b. City standardization. In the initial database, customer addresses were stated in an unstandardized format. Therefore, the address is standardized by only considering the city's name.

c. Removing the name column because, in this research, the phone number is the unique identifier of the consumer.

d. Standardize the month from Indonesian to English.

e. Delete rows that have a phone number value of 0 (NaN).

4. Modeling

This research focuses on analyzing loyal customers and customer segmentation. As mentioned in the previous step, the phone number will be used as a unique identifier because there was no unique identifier for each customer in the current database. In the modeling step, several techniques were used: descriptive statistics, retention analysis, Pareto charts, and customer segmentation.

a) Descriptive statistic

The descriptive statistics aimed to recognize the transaction pattern based on month, day, and city. Figure 4 depicts the total monthly transactions (Figure 4.a) and average transaction value (Figure 4.b). Total monthly transactions experienced a significant decline from May until November. Although total transactions decreased, average transactions increased significantly after April. This finding indicates that consumer spending has changed from buying in small quantities to buying in large quantities.

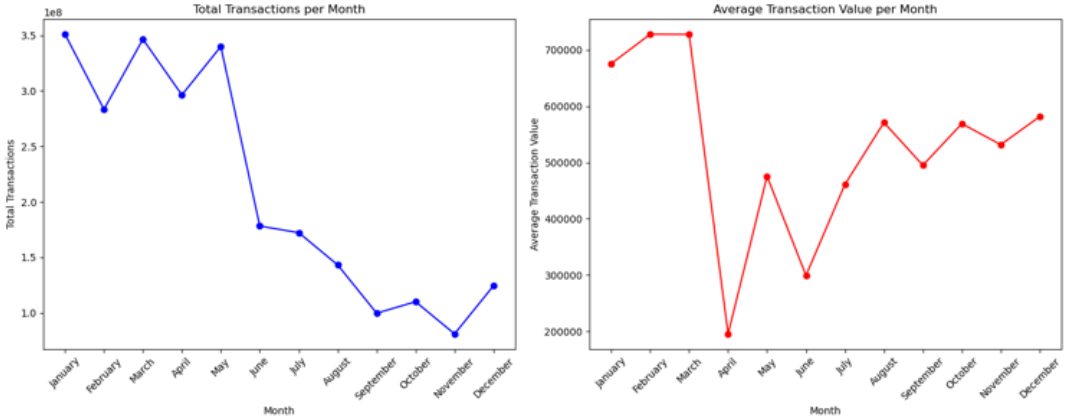


Figure 4. (a) Total monthly transaction, and (b) Average monthly transaction

Figure 5 indicates the daily transactions across all months. Most transactions occurred in the last week (fourth week in Figure 1), while the lowest occurred in the third week. Transactions increase significantly from the 23rd to the 27th. This period is the payroll period for most workers in Indonesia. Thus, companies can implement promotional policies to capture the increase in customer orders in the last week of each month.

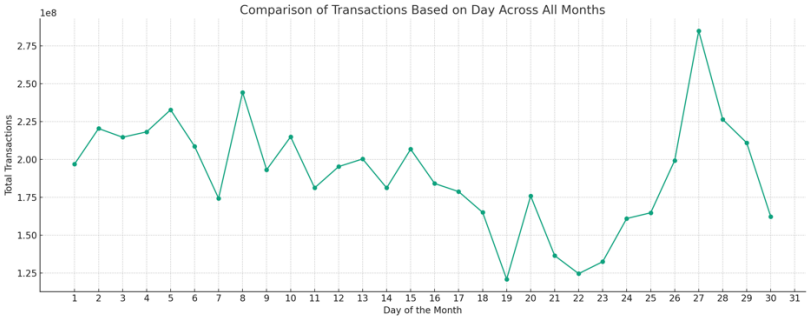


Figure 5. Comparison of Transaction Based on Day Across All Months

Figure 6.a indicates the number of transactions based on city, whereas Figure 6.b depicts the number of unique customers. Sleman, Yogyakarta, Kendal, Bekasi, and South Jakarta are the cities with the most transactions. The production location of the company used in this case study is in the Sleman area, so Sleman and Yogyakarta, the cities with the highest transactions and number of unique consumers, are reasonable. However, the high number of transactions and unique consumers in Kendal, Bekasi, and South Jakarta are interesting things to pay attention to. The company can take the free shipping policy to increase the number of transactions in cities other than DIY Province.

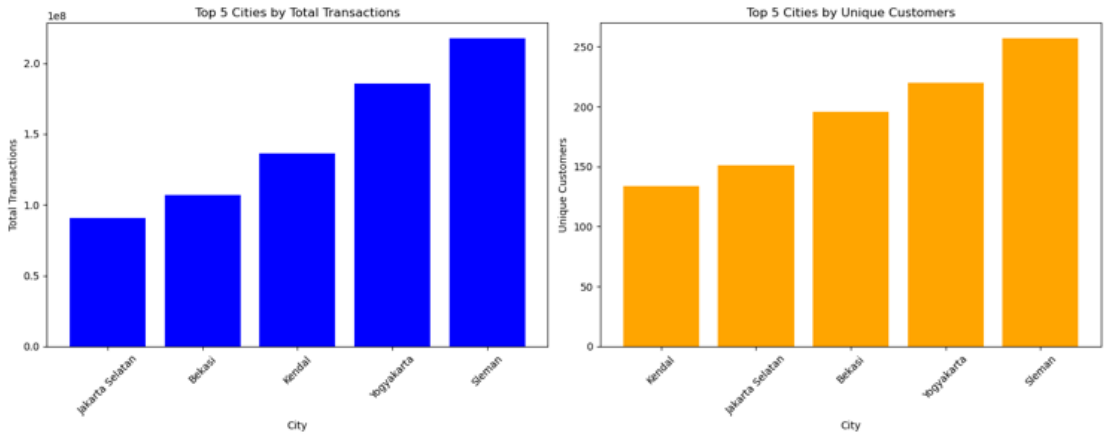


Figure 6. Geographical analysis: (a) top 5 cities by total transactions, and (b) top 5 cities by unique customers

b) Retention analysis

Figure 6 illustrates the customer retention rates. The retention rate is calculated as the percentage of customers from the previous month who remain customers in the current month. The highest retention occurs in April, July, and December. This finding indicates that of customers who purchased in March, 14% will buy in April. If reviewed more deeply, April is the month of fasting for Muslims. In addition, the increase also occurred in December, when this month is also a big day for Christians. Therefore, the recommendation that can be given to companies is to prepare attractive promotions during the month of religious holidays.

Figure 7. Customer retention rates and customer churn rates

c) Pareto analysis

The dataset consists of 3,271 unique customers. The Pareto analysis obtained 1,366 customers, contributing to 80% of the total transaction value (Figure 7). These 1,366 customers represent 41,76% of the total customer. This result is slightly different from the 80:20 classic Pareto Principle, which states that 80% of total value is contributed by 20% of customers.

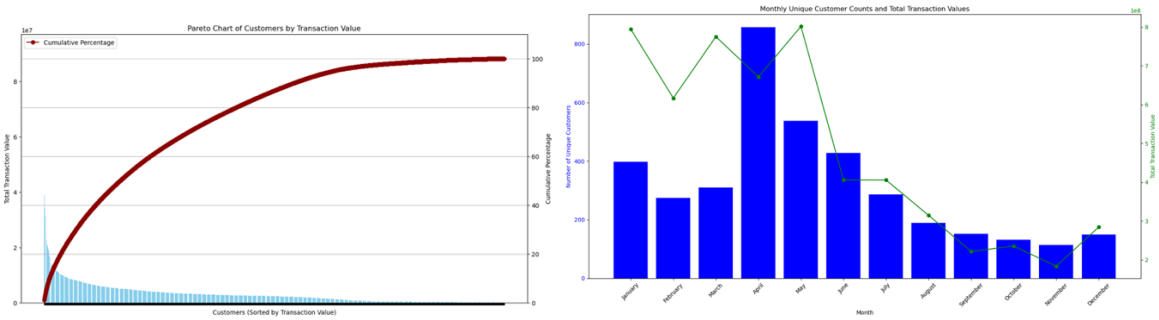


Figure 8. (a) Pareto chart and (b) Monthly unique customer

Approximately 31.29% of the total transaction value is attributed to loyal customers. This study defines a loyal customer as a customer who has made more than one transaction across the dataset. On the other hand, around 68.71% of the total transaction value comes from new customers, defined as those who have made only one transaction. Figure 8 further displayed monthly unique customers. The number of unique customers increased in April and December, whereas in the previous section, this increase was caused by religious events. This analysis indicates that new customers generate a significant portion of transaction value. However, loyal customers, though fewer in number, also contribute a substantial part of the total transaction value. Therefore, the recommendation given to the observed company is (1) increasing the number of new customers, (2) converting new customers to loyal customers, and maintaining the number of loyal customers.

d) Customer Segmentation

Figure 8 shows that the distribution of most consumers made 1 (one) transaction, with a value ranging below IDR 1 million. This finding indicates that most consumers conduct transactions in low quantities and values.

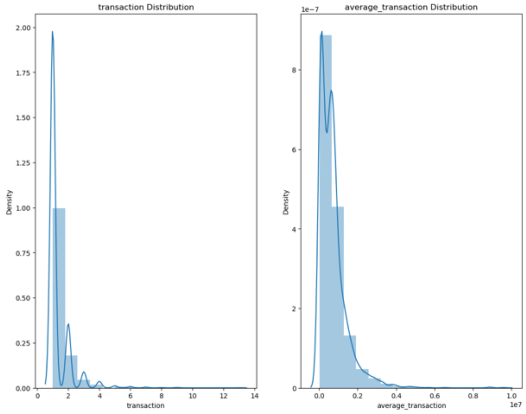


Figure 9. Density analysis on “transaction” and “average-transaction”

As a reference for determining the number of clusters, the elbow method is used in this study to determine the best number of clusters. The results of the elbow method obtained the best number of clusters, which is 3 clusters (Figure 9.a).

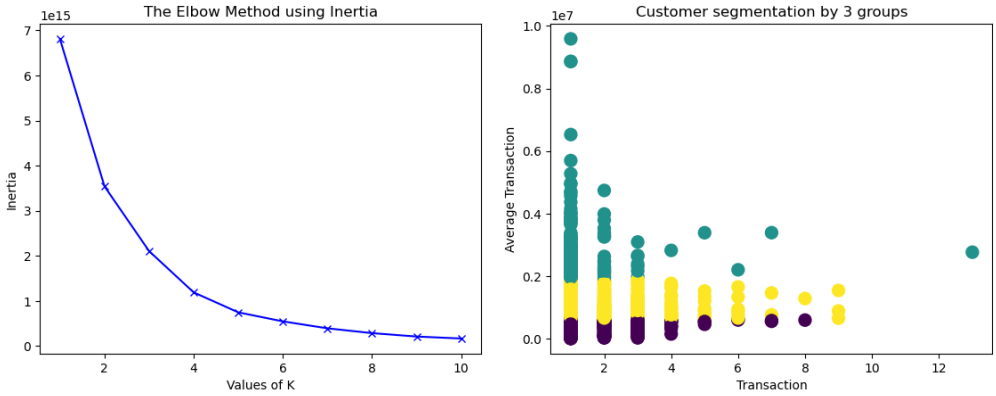


Figure 10. (a) Elbow method and (b) customer segmentation

The K means clustering results obtained 3 clusters illustrated in the scatter plot as shown in Figure 9b.

Cluster 1. High number of transactions, high average value of transactions: 192 customers.
Cluster 2. Medium number of transactions, medium average value of transaction: 1264 customers.
Cluster 3. Low number of transactions, low average value of transaction: 1815 customers.

Further analysis is carried out on cluster 1 to see customers' behavior included in the high transaction category and transaction value. The buying behavior of two customers from cluster 1 is depicted in Figure 10. Consumer number 82136370755 tends to make large purchases in May, where the results of the previous analysis of the month coincided with the religious day of Eid al-Fitr. Meanwhile, consumer number 85290006888 tends to make large purchases in December, coinciding with the religious day of Christmas.

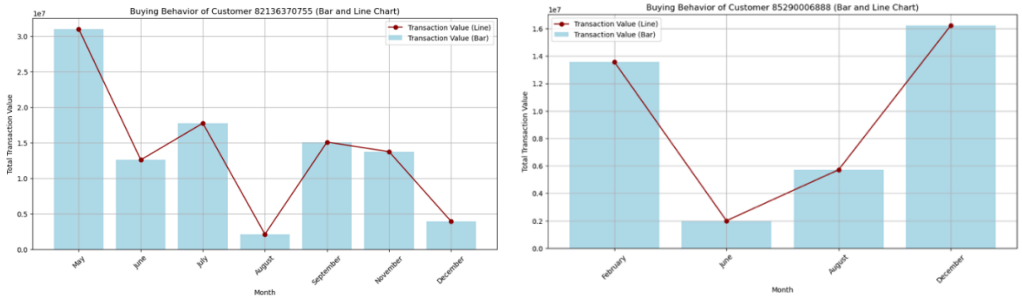


Figure 11. Buying pattern of two customers in Cluster 1:
(a) customer 82136370755, and (b) customer 85290006888

5. Evaluation

Based on the analysis carried out in the previous stage, the evaluation stage recapitulates the findings from the analysis. These findings are then used to provide recommendations for companies to improve their competitiveness.

a. Consumer spending has changed from buying in small quantities to buying in large quantities.
b. Most transactions occurred in the last week, while the lowest occurred in the third week.
c. Apart from Yogyakarta, which is the location of production and display stores, Jakarta and Bekasi are the cities with the largest unique consumers and sales.
d. Sales increase during religious events (Idul Fitri and Christmas).
e. 41.76% of the total customers contribute to 80% of sales.
f. 31.29% of the total transaction value is attributed to loyal customers.
g. 192 customers are in cluster 1 (high number of transactions and high transaction value).

Recommendations are given based on the findings of the above analysis and discussions with the company's management team. The recommendations given are as follows:

a. Provide discounted prices for customers who buy in large quantities.
b. Provide promotions and increase production capacity in the 4th week of each month and on religious days (Eid al-Fitr and Christmas).
c. Provide discounted shipping promotions for areas outside Jogja.
d. Conduct programs to reach loyal customers and turn new customers into loyal customers.

3.2. Discussion

This research focused on data mining to analyze sales data in a company that produces traditional motif clothing- batik. Data mining techniques in this study were implemented through 5 steps: business understanding, data understanding, data preparation, modeling, and evaluation. In the modeling stage, several analysis techniques are used to evaluate consumer behavior: descriptive statistics, retention analysis, Pareto charts, and customer segmentation. The evaluation step obtained several findings related to customer behavior, as presented in the evaluation stage.

Compared to previous research, this finding has similarities with previous research. Previous research found that the pandemic has led to a change in consumer preference, with a preference for online shopping, including for clothing [17]. Previous research has also found that after the pandemic, consumers tend to refrain from buying clothes and prioritize basic needs, which impacts the decline in clothing sales [18]. Using observation and interview methods, Khan and Sharma [19] found a positive relationship between consumer purchases and religious holidays. This is relevant to the findings on batik sales that increase in the month coincides with religious holidays.

This research can contribute to both managerial implications and theoretical implications. In managerial implications, the findings in this study can be the basis for the observed companies in particular and all batik producers to formulate marketing strategies. Regarding data mining procedures, this research uses a phone number as a unique identifier for each customer. Therefore, each customer should use a unique identifier, such as a customer number, to facilitate customer identification. On the other hand, this research can also provide theoretical implications. In the results of previous reviews, no research comprehensively uses data mining to analyze sales in batik producers. Research by Mulyawan et al. [12] studied batik sales in terms of motifs and sales using clustering techniques. Shop owners then used the results to formulate sales strategies. On the other hand, research by Bhaskara et al. [20] focused on creating a data warehouse system to record and monitor sales transaction data for batik companies. The researcher stated that real-time access to information could help companies develop their strategies. In addition, the data warehouse is the first step in the data mining process in this study.

4. Conclusion

The analysis using five data mining stages obtained insights that can be used as a basis for formulating batik sales strategies. This research focuses on analyzing customer purchasing behavior and evaluating loyal consumers. The analysis obtained insights into consumer behavior: changes from purchases in small quantities to large quantities, increased purchases in the final week of each month, and increased purchases on religious days. In addition, the analysis obtained that 31.29% of total transactions were contributed by loyal consumers, and 192 consumers were in the cluster of high transaction amounts and high transaction values. Future research can be performed by applying data mining to analyze the types of products the company sells. The results of discussions with management indicate that changes in consumer shopping behavior also occur in the types of products and batik motifs chosen by consumers.

References

[1] J. Bermudez, K. Apolinario, and A. G. Abad, “Layout Optimization and Promotional Strategies Design in a Retail Store based on a Market Basket Analysis,” 2016.

[2] M. Kaur and S. Kang, “Market Basket Analysis: Identify the Changing Trends of Market Data Using Association Rule Mining,” *Procedia Comput. Sci.*, vol. 85, no. Cms, pp. 78–85, 2016, doi: 10.1016/j.procs.2016.05.180.

[3] J. Kenyon and M. Vakola, “Customer Relationship Management: A Viable Strategy for the Retail Industry?,” *Int. J. Organ. Theory Behav.*, vol. 6, no. 3, pp. 329–353, 2003, doi:

https://doi.org/10.1108/IJOTB-06-03-2003-B001.
[4] R. Karthiyayini and R. Balasubramanian, “Affinity and Association Affinity Analysis and Association Rule Mining using Apriori Algorithm in Market Basket Analysis,” *Int. J. Adv. Res. Comput. Sci. Softw. Eng.*, vol. 6, no. 10, pp. 241–246, 2016, [Online]. Available: www.ijarcsse.com.
[5] M. Samuel and Z. Gwendolyn, “An Investigation of the Role of Product , Place , Promotion and Price (4Ps) in Market Basket Analysis : A Case Study of Retail Shops in Gweru Zimbabwe,” *Int. J. Sci. Res.*, vol. 3, no. 2, pp. 132–135, 2014.
[6] S. Panjaitan *et al.*, “Implementation of Apriori Algorithm for Analysis of Consumer Purchase Patterns,” *J. Phys. Conf. Ser.*, vol. 1255, no. 1, 2019, doi: 10.1088/1742-6596/1255/1/012057.
[7] A. K. S. Ong *et al.*, “Consumer behavior in clothing industry and its relationship with open innovation dynamics during the covid-19 pandemic,” *J. Open Innov. Technol. Mark. Complex.*, vol. 7, no. 4, p. 211, 2021, doi: 10.3390/joitmc7040211.
[8] T. N. Hashem, “Examining the Influence of COVID 19 Pandemic in Changing Customers' Orientation towards E-Shopping,” *Mod. Appl. Sci.*, vol. 14, no. 8, p. 59, 2020, doi: 10.5539/mas.v14n8p59.
[9] L. Eger, L. Komárková, D. Egerová, and M. Mičík, “The effect of COVID-19 on consumer shopping behaviour: Generational cohort perspective,” *J. Retail. Consum. Serv.*, vol. 61, no. February, 2021, doi: 10.1016/j.jretconser.2021.102542.
[10] R. A. Rahadi, D. Rahmawati, N. A. Windasari, and P. F. Belgiawan, “The Analysis of Consumers ’ Preferences for Batik Products in Indonesia,” *Rev. Integr. Bus. Econ. Res.*, vol. 9, no. 1, pp. 278–287, 2020.
[11] R. Ardiyansyah and R. A. M. Febrianti, “Understanding the driver of customer purchase decision : The role of customer engagement and brand attachment in Batik Products,” *Fair Value J. Ilm. Akunt. dan Keuang.*, vol. 5, no. 4, pp. 1979–1985, 2022.
[12] Mulyawan, A. Bahtiar, and I. Ali, “Transformasi Strategi Penjualan Batik Cirebon Dengan Pendekatan Analisis Pengelompokan K-Means,” *KOPERTIP J. Ilm. Manaj. Inform. dan Komput.*, vol. 7, no. 1, pp. 1–7, 2023, doi: 10.32485/kopertip.v7i1.309.
[13] V. Salbinda, R. I. Handayani, and N. Normah, “Clustering of Clothing Sales Data At Top Store Using K-Means Method,” *SinkrOn*, vol. 7, no. 2, pp. 314–320, 2022, doi: 10.33395/sinkron.v7i2.11244.
[14] M. Mardalius and T. Christy, “Mapping of Potential Customers As a Clothing Promotion Strategy Using K-Means Clustering Algorithm,” *... J. Ilmu Pengetah. Dan ...*, vol. 6, no. 1, pp. 67–72, 2020, doi: 10.33480/jitk.v6i1.1414.
[15] A. Anshu, “Review Paper on Data Mining TechniquesandApplications,” *Int. J. Innov. Res. Comput. Sci. Technol.*, vol. 7, no. 2, pp. 22–26, 2019, doi: 10.21276/ijircst.2019.7.2.4.
[16] M. A. Jassim and S. N. Abdulwahid, “Data Mining preparation: Process, Techniques and Major Issues in Data Analysis,” *IOP Conf. Ser. Mater. Sci. Eng.*, vol. 1090, no. 1, p. 012053, 2021, doi: 10.1088/1757-899x/1090/1/012053.
[17] H. Shen, F. Namdarpour, and J. Lin, “Investigation of online grocery shopping and delivery preference before, during, and after COVID-19,” *Transp. Res. Interdiscip. Perspect.*, vol. 14, no. March, p. 100580, 2022, doi: 10.1016/j.trip.2022.100580.
[18] C. Liu, S. Xia, and C. Lang, “Clothing Consumption During the COVID-19 Pandemic: Evidence From Mining Tweets,” *Cloth. Text. Res. J.*, vol. 39, no. 4, pp. 314–330, 2021, doi: 10.1177/0887302X211014973.
[19] M. R. Khan and K. Sharma, “Purchase Preferences and Buying Influences on Religious Occasions,” *FIIB Bus. Rev.*, vol. 9, no. 3, pp. 216–227, 2020, doi: 10.1177/2319714520942624.
[20] I. M. A. Bhaskara, L. G. P. Suardani, and M. Sudarma, “Data Warehouse Implemantation To Support Batik Sales Information Using MOLAP,” *Int. J. Eng. Emerg. Technol.*, vol. 3, no. 1, pp. 45–51, 2018.

Advance Sustainable Science, Engineering and Technology (ASSET)
Vol. 6, No.2, April 2024, pp. 02402016-01 ~ 02402016-08
ISSN: 2715-4211 DOI: https://doi.org/10.26877/asset.v6i2.18360

Meta Insights: Analyzing Hero Performance in Mobile Legends with K-Nearest Neighbors

Octarizal Kristanto*, Pulung Nurtantio Andono.

Faculty of Computer Science, Universitas Dian Nuswantoro, Jl. Imam Bonjol No.207, Semarang, Central Java 50131, Indonesia

*111202012635@mhs.dinus.ac.id

Abstract. This research presents a thorough statistical analysis of hero performance in the latest Mobile Legends meta, employing the K-Nearest Neighbor (KNN) algorithm. Utilizing diverse data sources, the study explores factors influencing hero success, leveraging KNN's ability to identify intricate patterns in complex datasets. Through meticulous data collection, preprocessing, and application of the KNN algorithm, the research classifies and predicts hero performance based on similarities with neighboring heroes. Critical determinants such as win rates, popularity, and hero ban rates emerge, providing profound insights into gameplay strategies. The study emphasizes the importance of understanding meta dynamics, hero attributes, and player expertise for informed decision-making in hero selection within the dynamic landscape of Mobile Legends.

Keywords: Mobile Legends; K-Nearest Neighbors; Hero Performance; Latest Meta; Statistical Analysis; Gameplay Strategies

(Received 2024-02-05, Accepted 2024-03-19, Available Online by 2024-03-19)

1. Introduction

In the ever-changing landscape of online battle arena (MOBA) gaming, Mobile Legends has emerged as a prominent game, demanding players to showcase tactical skills and strategic thinking in hero selection and utilization. Maintaining an edge in the evolving meta is crucial for strategic advantage in Mobile Legends, given the dynamic nature of the game that requires constant adaptation. This research aims to contribute significantly to the field by conducting an in-depth statistical analysis of hero performance in Mobile Legends, with a specific focus on the latest meta. To achieve this objective, the study will leverage the K-Nearest Neighbor (KNN) algorithm due to its ability to classify data based on similarity, making it a potent tool for identifying complex patterns and relationships within the extensive dataset of hero statistics.[1] Aligned with the multidimensional nature of hero performance metrics in Mobile Legends, including win rates, pick rates, and kill-to-death ratios, KNN excels in identifying patterns by considering the proximity of data points in this multidimensional space. This approach promises a nuanced understanding of hero interactions and effectiveness within the current meta, surpassing conventional observations. [2] The research not only aims to uncover trends but also intends to explore the implications of these findings on strategic decision-making in Mobile Legends.

By harnessing the potential of the KNN algorithm, the analysis seeks to offer a profound understanding of the underlying dynamics that influence hero success. [3] The methodology involves the collection and analysis of extensive datasets, capturing the nuances of hero performance in diverse in-game scenarios. Essentially, the combination of advanced statistical analysis and the KNN algorithm promises to illuminate the intricate tapestry of hero performance, providing a valuable resource for players, analysts, and enthusiasts navigating the constantly evolving landscape of Mobile Legends.[17] Additionally, this section will discuss specific techniques used to demonstrate the performance of the KNN algorithm, providing a thorough explanation of why KNN was chosen and offering a more detailed insight into how this algorithm is implemented in the context of this research. This aims to provide a comprehensive and in-depth overview for readers regarding the contribution and utility of KNN in analyzing hero performance in Mobile Legends.

2. Methods

2.1. Research Stages

The research on hero performance data in Mobile Legends was conducted through a systematic and comprehensive process outlined in Figure 1, illustrating the flowchart of the data analysis process.

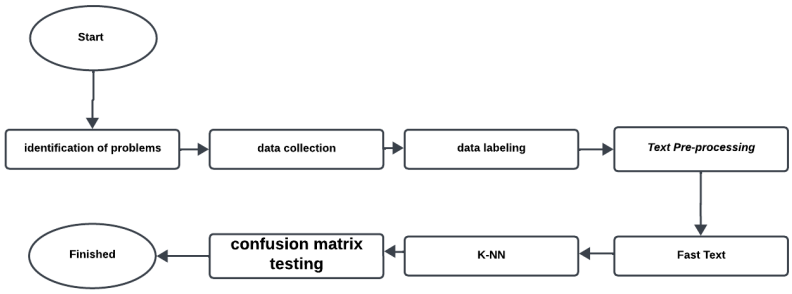


Figure 1. Flowchart. Data Analysis Process

The initiation of the research involved an exhaustive collection of hero performance data from diverse and reliable sources, including official game repositories, player databases, and esteemed community-driven platforms specializing in sharing in-game statistics. [9] This comprehensive approach aimed to gather a wealth of attributes and relevant statistics for Mobile Legends heroes, forming the basis for subsequent analyses and ensuring a holistic understanding of hero performance.

Following data collection, the collected dataset underwent a rigorous preprocessing phase to validate and ensure its integrity, mitigating potential inconsistencies or errors. This critical step involved meticulous preparation of the data, establishing a robust foundation for meaningful analysis.

2.2. Algoritma K-Nearest Neighbors (KNN)

The KNN method was applied to analyze hero performance. This approach allows us to classify and predict performance based on similarity with other heroes in the dataset.[9] K-Nearest Neighbors algorithm, a non-parametric and instance-based learning method, was chosen for its capability to make predictions based on the proximity of instances in a multidimensional feature space. This method operates on the principle that instances with similar features are likely to share similar outcomes.[2]

To implement the KNN algorithm for hero performance analysis, a thorough examination of the dataset was conducted to identify relevant features and establish a meaningful similarity metric. Subsequently, the algorithm was applied to classify and predict the performance of each hero by considering the characteristics and attributes shared with its nearest neighbors.

The selection of an appropriate value for the 'k' parameter, representing the number of neighbors to consider, was a crucial aspect of the implementation. This value was determined through a systematic evaluation process, considering the dataset's characteristics and the desired level of granularity in performance analysis.
The outcomes of the KNN algorithm were then assessed and validated through rigorous testing procedures, ensuring the reliability and accuracy of the performance predictions. The results obtained from this implementation provided valuable insights into the relationships and patterns within the hero dataset, contributing to a more nuanced understanding of hero performance dynamics[19].

$$dis\,(x_1, x_2) = \sqrt{\sum_{i=0}^{n} (x_{1i} - x_{2i})^2}$$

Figure 2. K-Nearest Neighbors Formula

The distance d between instances Xi and Xj in a feature space is often calculated using the Euclidean distance formula:

$$d(Xi, Xj) = \sum l = 1n(Xil - Xjl)2$$

where n is the number of attributes.

The classification probability for an instance Iij belonging to class Cci is computed as follows:

$P(I=ij|C=ci)=k1\sum l=1kp(I=ij|C=cl)$

Explanation:

- $p(I=ij|C=cl)$: Probability of instance ij given class cl based on the k-nearest neighbors.
- $p(C=cl)$: Prior probability of class cl in the dataset.

In the context of predicting hero characteristics in MPL Season 12, the KNN algorithm can be applied to features such as Hero Type, Hp Hero, Mana Hero, Defend Hero, and Attack Hero. The algorithm predicts the characteristics of a hero based on the attributes of its k-nearest neighbors in the dataset, considering factors like weapon utilization, health power, mana usage, defense, and attack statistics. The choice of k influences the granularity of the predictions, with smaller k values providing more local predictions and larger k values offering a broader perspective.

3. Results and Discussion

Embarking on a thorough examination of our study's methodologies, we delve deep into the nuanced analysis of hero performance within the dynamic Mobile Legends meta. At the heart of our investigative approach lies the strategic utilization of the K-Nearest Neighbors (KNN) algorithm, renowned for its adeptness in deciphering intricate patterns within multidimensional datasets.

The primary objective of our study is to conduct a comprehensive statistical analysis of hero performance in Mobile Legends, with a keen focus on unraveling the subtleties of the latest meta. The deliberate choice of the KNN algorithm is grounded in its non-parametric and instance-based learning capabilities, making it particularly well-suited for predicting outcomes based on the proximity of instances within a multidimensional feature space. The rationale behind opting for KNN lies in its efficacy in analyzing the multifaceted nature of hero performance metrics in the dynamic Mobile Legends landscape. The algorithm operates on the principle that instances with similar features are likely to exhibit comparable outcomes, proving invaluable in a context where hero performance is intricately influenced by factors such as win rates, pick rates, and kill-to-death ratios. Thus, the KNN algorithm becomes instrumental in unraveling complex patterns, providing nuanced insights into hero interactions

and effectiveness within the current meta. [12]. The core formula guiding the KNN algorithm for predicting outcomes is succinctly expressed as:

$$Y=\text{mode}(Yi)$$

where:

- Y signifies the predicted outcome for a given hero.
- Yi represents the outcomes of the k nearest neighbors.

The essence of the KNN algorithm lies in predicting outcomes based on the majority class of nearest neighbors. Our research journey begins with the meticulous collection of hero performance data from diverse and reliable sources, including community data-sharing platforms. This dataset, comprising various attributes like win rates, popularity, HP, attack speed, and magical power, forms the foundation for our analysis. Before applying the KNN algorithm, we undergo a critical preprocessing phase to ensure dataset integrity, eliminating inconsistencies or anomalies. With the dataset curated, we employ the KNN algorithm to unveil hero performance intricacies in Mobile Legends, considering factors like win rates and specific hero characteristics. The algorithm's effectiveness is evident in uncovering complex patterns, providing nuanced insights into hero interactions within the dynamic meta. Our study contributes to mobile gaming's evolving landscape by enhancing understanding of hero attributes and in-game performance dynamics. This empowers players, aids analysts in strategy formulation, and offers valuable insights to game developers for refining the gaming experience. In essence, our research enriches the dialogue on mobile gaming dynamics, augmenting collective knowledge within the gaming community.

HERO	WIN	POPULARITY	BANNED	HP	Physical Attack	Physical Defense	Attack Speed
Rafaela	56.94%	1.33%	15.38%	2571	124	18	1
Ixia	55.14%	1.10%	5.83%	2620	113	17	1.3
Mathilda	54.97%	0.47%	66.24%	2651	120	22	1
Lolita	54.80%	0.57%	31.75%	3051	115	30	0.98
Terizla	54.70%	1.51%	2.08%	2728	159	25	1
Irithel	54.59%	1.96%	1.25%	2540	132	17	1.23
Yi Sun-shin	54.49%	0.25%	0.06%	2570	114	22	1.03
Baxia	54.15%	0.43%	0.62%	3369	125	39	1
Edith	54.05%	1.57%	3.44%	3149	110	27	1.04
Argus	54.03%	0.23%	0.17%	2628	118	21	1.19
Minotaur	53.60%	0.52%	0.46%	3459	123	34	0.91
Belerick	53.42%	0.64%	0.63%	3641	110	25	1.01
Benedetta	53.40%	0.53%	0.26%	2569	131	18	1.11
Hylos	53.40%	0.28%	0.09%	3309	105	33	1.03

Table 1. Dataset Hero

Critical Chance	Mana	Magic Power	Magical Defenses	Cooldown Reduction	Movement SPD	MYTHIC+
0%	680	8	17	10%	330	HIGH
0%	440	0	15	0%	300	LOW
0%	430	38	15	5%	325	HIGH
0%	480	0	25	0%	335	HIGH
0%	430	0	21	0%	318	HIGH
0%	0	0	15	0%	325	HIGH

5%	438	0	15	0%	320	LOW
0%	422	0	31	0%	300	HIGH
0%	0	0	25	0%	318	HIGH
0%	0	0	15	0%	332	LOW
0%	0	0	26	0%	325	HIGH
0%	450	0	25	0%	312	LOW
0%	460	0	15	0%	326	HIGH
0%	430	0	31	0%	325	LOW
0%	550	52	15	2%	320	LOW

Table 2. Dataset Hero

The effectiveness of any predictive model, such as KNN, relies on its capacity to deliver precise and dependable outcomes. Through rigorous validation and refinement against actual results, the algorithm ensures close alignment with real-world scenarios. This iterative process not only enhances prediction accuracy but also strengthens the credibility of analytical insights derived from the KNN algorithm. The emphasis is placed on uncovering complex patterns and attaining nuanced insights into the multifaceted dynamics of hero interactions within the latest meta.[14]

HP	Physical Attack	Physical Def	Attack Speed	Critical Chance	Mana	Magic Power	Magical Def	Cooldown Reduce	Movement Speed
2571	124	18	1	0	680	8	17	0.1	330
2628	118	21	1.19	0	0	0	15	0	332
2778	139	29	1.05	0	0	0	21	0	312
2801	128	22	1.04	0.05	0	0	15	0	332
2573	114	21	1	0	493	38	15	0.05	309
2698	159	27	1	0	0	0	21	0	306
2490	123	20	1.21	0	435	0	15	0	333
2501	113	17	1.01	0	500	38	15	0.05	316

Table 3. ExampleSet

Row No.	MYTHIC+	prediction(MYTHIC+)	prediction(HIGH)	confidence(LOW)	Hero	Win	Popularity	Banned
1	HIGH	LOW	0.433306419	0.566693581	Rafaela	0.5694	0.0133	0.1538
2	LOW	HIGH	0.770063276	0.229936724	Argus	0.5403	0.0023	0.0017
3	HIGH	HIGH	0.805117389	0.194882611	Khaleed	0.5308	0.0018	4.00E-04
4	HIGH	HIGH	0.804136143	0.195863857	Freya	0.5227	0.004	0.003
5	LOW	LOW	0.202542328	0.797457672	Alice	0.5151	0.0012	7.00E-04
6	HIGH	HIGH	0.799617267	0.200382733	Yu Zhong	0.5127	0.0125	0.0029
7	HIGH	LOW	0.381287917	0.618712083	Brody	0.5113	0.0103	0.0031
8	HIGH	HIGH	0.608041496	0.391958504	Lylia	0.5111	0.0071	0.0024

Table 4. ExampleSet

K-Nearest Neighbors (KNN) is a machine learning algorithm utilized to forecast outcomes by assessing the similarity among instances. In our study, it's employed to anticipate the rank of heroes (MYTHIC+, HIGH, LOW) utilizing attributes outlined in the ExampleSet table.

1. Euclidean Distance:

- Compute the Euclidean distance between the hero under scrutiny and all other heroes within the dataset.
- Distance=$\sum j=1n(Xij-Xkj)2$, where Xij and Xkj represent attribute values for feature j of the hero under scrutiny and hero k, respectively, and n denotes the number of features.

2. Identifying Nearest Neighbors:
- Choose the k nearest neighbors based on the smallest distances.

3. Determining Rank Prediction:
- Forecast the rank of the hero under scrutiny based on the predominant rank among its neighbors.

accuracy: 61.54%			
	true HIGH	**true LOW**	**class precision**
pred. HIGH	**4**	**2**	**66.67%**
pred. LOW	**3**	**4**	**57.14%**
class recall	**57.14%**	**66.67%**	

Table 5. Performance Vector

The application of the K-Nearest Neighbors (KNN) algorithm in our study has provided valuable insights into hero performance within the dynamic Mobile Legends meta. Let's explore these insights by focusing on the key aspects derived from our analysis: [15]

1. **Outcome Prediction:**
- The KNN algorithm predicts outcomes based on the majority class of the k nearest neighbors.
- Hero performance data are meticulously collected and preprocessed to facilitate accurate predictions.

2. **Validation and Refinement:**
- Accuracy and reliability of predictions are ensured through rigorous validation against real-world outcomes.
- Iterative validation enhances prediction accuracy and reinforces analytical credibility.

3. **Performance Metrics:**
- Predictions for hero attributes, including MYTHIC+ rank, confidence levels, and corresponding hero information, are presented.
- Probability of winning and overall accuracy of the KNN algorithm are visualized.

4. **Effectiveness of KNN Algorithm:**
- The algorithm, based on the proximity principle, predicts outcomes by considering the similarity of hero attributes.
- Nuanced insights into hero interactions within the dynamic Mobile Legends meta are provided.
- Uncovering complex patterns contributes to actionable insights for players, analysts, and game developers.

In conclusion, the KNN algorithm effectively predicts hero performance, enhancing understanding of the Mobile Legends meta. Its iterative validation process ensures practical utility in gaming contexts, providing accurate predictions aligned with real-world scenarios.

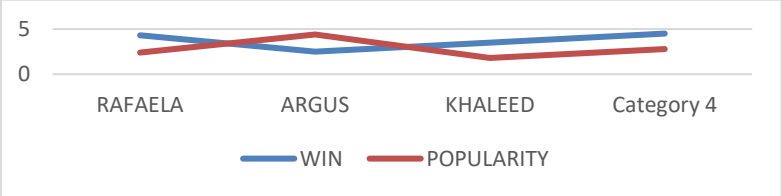


Figure 3. Probability of Winning

Figure 3 visualizes the probability of winning based on the predictions made by the K-Nearest Neighbors (KNN) algorithm. The x-axis represents different instances, while the y-axis shows the corresponding probability of winning. This graph provides an overview of the algorithm's predictive performance in estimating the success of heroes within the Mobile Legends meta.

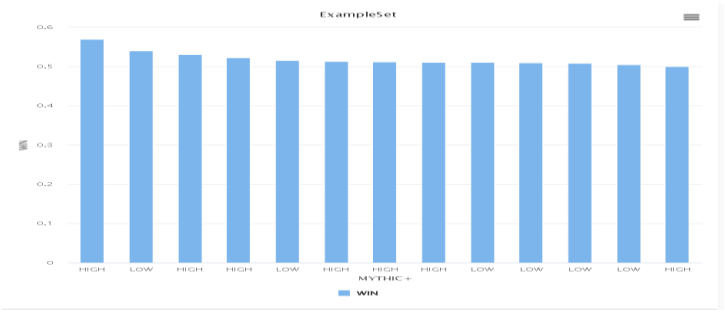


Figure 4. Accuration

Figure 4 depicts the comprehensive accuracy of the KNN algorithm in forecasting hero attributes. The x-axis denotes distinct iterations or instances, while the y-axis showcases the accuracy percentage. This graphical representation showcases the algorithm's efficacy in predicting hero ranks (MYTHIC+, HIGH, LOW) using the given attributes. Accuracy serves as a pivotal metric in evaluating the predictive model's dependability.

4. Conclusion

In conclusion, this research aims to conduct a comprehensive statistical analysis of hero performance in the latest Mobile Legends meta using the K-Nearest Neighbor Algorithm (KNN). The study provides valuable insights into the various factors influencing hero success. The methodology involves a rigorous research stage, encompassing the collection of hero performance data from diverse sources and community data-sharing platforms. The dataset undergoes detailed pre-processing to ensure integrity and validity, establishing a solid foundation for subsequent applications of the KNN algorithm. The selected algorithm, renowned for its non-parametric and instance-based learning capabilities, effectively uncovers complex patterns in the multidimensional space of hero statistics.

The research results present a comprehensive statistical profile of hero performance, demonstrated through the application of the KNN algorithm. This analysis reveals crucial factors, including win rates, popularity, and hero ban rates, offering different insights into interactions in the latest Mobile Legends meta. The study outlines performance vectors, accuracy metrics, and factors influencing hero performance. Key factors affecting a hero's success encompass win rate, popularity, ban rate, hero characteristics (hero kit), meta adaptation, and player skill. The study underscores the importance of a deep understanding of the meta game, hero abilities, and effective gameplay strategies for proper hero selection. While this research provides valuable insights for Mobile Legends enthusiasts, analysts, and industry professionals, it also acknowledges certain limitations. Future research efforts could address these limitations, refining the methodology used in this study. In essence, this research serves as a valuable resource, offering a detailed exploration of the dynamics of hero performance in the ever-evolving Mobile Legends landscape.

References

[1] M. Z. Mustamiin, “KONSELING INDIVIDU DENGAN SIKAP KECANDUAN GAME ONLINE MOBILE LEGEND PADA SISWA.”

[2] M. Stmik and N. Mandiri, “Penerapan Algoritma K-Means Clustering pada Karakter Permainan Multiplayer Online Battle Arena,” *JURNAL INFORMATIKA*, vol. 6, no. 2, pp. 246–254, 2019, [Online]. Available: http://ejournal.bsi.ac.id/ejurnal/index.php/ji

[3] W. Wang and R. Bierig, “Predicting Multiplayer Online Battle Arena (MOBA) Game Outcome Based on Hero Draft Data MSc Research Project Data Analytics”, [Online]. Available: https://steamdb.info/app/570/graphs/
[4] S. Zhang, D. Cheng, Z. Deng, M. Zong, and X. Deng, “A novel kNN algorithm with data-driven k parameter computation,” *Pattern Recognit Lett*, vol. 109, pp. 44–54, Jul. 2018, doi: 10.1016/j.patrec.2017.09.036.
[5] M. Mora-Cantallops and M. Á. Sicilia, “MOBA games: A literature review,” *Entertain Comput*, vol. 26, pp. 128–138, May 2018, doi: 10.1016/j.entcom.2018.02.005.
[6] I. G. W. Sena and A. W. R. Emanuel, “MOBILE LEGEND GAME PREDICTION USING MACHINE LEARNING REGRESSION METHOD,” *JURTEKSI (Jurnal Teknologi dan Sistem Informasi)*, vol. 9, no. 2, pp. 221–230, Mar. 2023, doi: 10.33330/jurteksi.v9i2.1866.
[7] T. An, “CROSS-CULTURE IMPACT ON VIDEO GAME INDUSTRY Case study: Dota 2’s game design and player experience.”
[8] C. S. Brunt, A. S. King, and J. T. King, “The influence of user-generated content on video game demand,” *Journal of Cultural Economics*, vol. 44, no. 1, pp. 35–56, Mar. 2020, doi: 10.1007/s10824-019-09349-0.
[9] A. S. Chan, F. Fachrizal, and A. R. Lubis, “Outcome Prediction Using Naïve Bayes Algorithm in the Selection of Role Hero Mobile Legend,” in *Journal of Physics: Conference Series*, Institute of Physics Publishing, Jul. 2020. doi: 10.1088/1742-6596/1566/1/012041.
[10] J. S. Raikwal and K. Saxena, “Performance Evaluation of SVM and K-Nearest Neighbor Algorithm over Medical Data set,” 2012.
[11] R. L. Zhu, J. Wang, and S. J. Zhao, “Categorical Learning-based Line-up Prediction in the Drafting Process of MOBA Games,” *Journal of Internet Technology*, vol. 24, no. 2, pp. 411–419, Mar. 2023, doi: 10.53106/160792642023032402019.
[12] S. Sahara, R. Agung Permana, U. Bina Sarana Informatika, and S. Antar Bangsa, “METODE KNN PADA SENTIMENT ANALISIS REVIEW PRODUK GAME ANDROID,” Online, 2022. [Online]. Available: www.zdnet.com
[13] S. M. Listijo, T. Purwani, S. T. Galih, and T. Hafidzin, “PREDIKSI KEMENANGAN DAN SUSUNAN TIM PADA GAME MOBILE LEGENDS BANG BANG MENGGUNAKAN ALGORITMA NAÏVE BAYES.”
[14] I. Fitriana and M. M. Rois, “Analysis of Metaphors Applied in Mobile Legends Game,” 2023.
[15] R. L. Zhu, J. Wang, and S. J. Zhao, “Categorical Learning-based Line-up Prediction in the Drafting Process of MOBA Games,” *Journal of Internet Technology*, vol. 24, no. 2, pp. 411–419, Mar. 2023, doi: 10.53106/160792642023032402019.
[16] D. Wicaksono Wiroto, N. Husain, H. Taan, and A. Malik Iskandar, “Cohesiveness as a Symbolic Power of Mobile Legends Online Game Community”, doi: 10.14710/jis.
[17] I. Ketut, S. Yogatama, A. Putra Kharisma, and L. Fanani, “Analisis Faktor-Faktor Yang Memengaruhi Minat Pemain Dalam Permainan MOBA (Studi Kasus: Mobile Legends: Bang-Bang!),” 2019. [Online]. Available: http://j-ptiik.ub.ac.id
[18] L. Hanke and L. Chaimowicz, “A Recommender System for Hero Line-Ups in MOBA Games.” [Online]. Available: www.aaai.org
[19] Z. Ren, A. Liu, Y. Liang, J. Peng, and J. Ma, “Efficient Meta Reinforcement Learning for Preference-based Fast Adaptation.”
[20] M. Suznjevic and M. Matijasevic, “Application Context Based Algorithm for Player Skill Evaluation in MOBA games.” [Online]. Available: http://forums.heroesofnewerth.com/showthread.php?562680-

Advance Sustainable Science, Engineering and Technology (ASSET)
Vol. 6, No.2, April 2024, pp. 02402017-01 ~ 02402017-08
ISSN: 2715-4211 DOI: https://doi.org/10.26877/asset.v6i2.18636

Project Delay Factor Analysis Using Fault Tree Analysis (FTA) Method and Rescheduling with Critical Path Method (CPM) in Ducting Manufacturing Projects at PT. FRA

Festo Renza Ardhana Putra*, Dzakiyah Widyaningrum

[1]Industrial Engineering, University of Muhammadiyah Gresik, Jl. Sumatera 101 GKB Randu Agung Gresik Telp. 031 39514141, Fax. 031 3952585

*festorenza32@gmail.com

Abstract. Ducting manufacturing project at PT. The FRA was planned to be completed in 14 days, but was implemented in 19 days and 5 days late. The method to find the cause of project delays is Fault Tree Analysis (FTA), then analyze rescheduling with the Critical Path Method (CPM). The purpose of this study is to identify factors of work delay, determine new steps for project contractors and reschedule with alternative working hours and labor additions. The results of the analysis using the FTA method obtained the main cause of delay in drawing engineering work of 0.0002592 and obtained 4 new steps for project contractors. The results of rescheduling using the CPM method obtained a percentage increase in productivity with the alternative of adding working for 4 hours with a duration of 25% or 6 days faster than the normal duration of 19 days to 13 days and an additional cost of 91.96%, which is Rp. 21,500,000. While the percentage of productivity increase with alternative labor additions with a duration of 38.75% or 5 days faster than the normal duration of 19 days to 14 days and additional costs of 178.57%, namely Rp. 31,200,000. So it can be concluded by using the FTA method the main cause is drawing engineering work and 4 new steps are obtained and using the CPM method alternative scheduling is obtained to increase overtime work hours more effectively and additional costs are not too large, so that the company does not experience delays for future ducting projects.
Keywords: Project delay, Fault Tree Analysis, Critical Path Method, Reschedule.

(Received 2024-03-14, Accepted 2024-03-29, Available Online by 2024-04-22)

1. Introduction

Project delays occur if the contractor does not complete the project according to the time schedule agreed in the contract, then the delay has the impact of increasing project costs, increasing market risk, increasing production delay time and project delays[1]. Project management aims to execute projects well to minimize problems during project implementation, so an accurate project management system is needed for effective and efficient scheduling[2]. PT. FRA is a project based steel fabrication manufacturing and service company that aims to improve the effectiveness and efficiency of the production process[3]. The company is engaged in steel manufacturing, heavy equipment operation and maintenance[4]. Each project has a different design and function to process raw materials that are assembled to be formed to produce new materials to have added functions and value. Problems of PT. FRA is a delay ducting line 149-152 project that the project owner of PT. MBI. Every project has a

planned schedule, but schedules and deliverables in the field show different project completion times. The ducting project was planned to be completed in 14 days, but it was carried out in 19 days and 5 days late. Because there are many delay factors in the project, it is necessary to analyze the Fault Tree Analysis (FTA) method to analyze the factors causing delays and the Critical Path Method (CPM) method to get rescheduling with alternative work hours and alternative labor additions.

Fault Tree Analysis (FTA) is an analysis method by displaying drawings and evaluating paths in a system to minimize costs[5]. Fault Tree Analysis is used to identify the root of the fault tree[6]. FTAs are also used to identify the causes of failures in visual diagrams and logic models[7]. Critical Path Method (CPM) is a project management method with a focus on the timing of work activities and defined critical path project duration[8]. The Critical Path Method details each activity, job activity, normal time and cost data for the network[9]. Based on the description above, researchers are interested in conducting research, the advantages of both methods are being able to find out the work and factors causing delays and determine new steps for project contractors and get rescheduling with alternatives to increase work hours and alternative labor additions in future projects.

2. Methods

The research methods in this study are qualitative and quantitative to obtain information about delays in projects, determine delay factors and make descriptive explanations about these factors can occur[10]. The subjects of this study analyzed the factors causing delays obtained from observations, interviews and questionnaire results distributed to respondents of planning production control (ppc) workers[11]. Fault Tree Analysis (FTA) method to process questionnaire data to determine the probability value to calculate the minimum cut set value. The ducting line 149-152 project at PT FRA is the object of research which contains time schedule data and reschedule data to obtain the schedule and project implementation presented in table 1 and table 2.

Table 1. Ducting Project Plan Schedule (Source PT. FRA)

NO	Job Description	Work Implementation Plan Schedule (Days)													
		1	2	3	4	5	6	7	8	9	10	11	12	13	14
I	Engineering drawing work														
II	Material procurement work														
A	Cutting marking production work														
B	Machining production work														
C	Feedup production work														
D	Welding production work							Holiday							Holiday
E	Finishing work														
F	Quality control work														
G	Painting work														
H	Packing delivery work														

Table 2. Reschedule Ducting Project (Source PT. FRA)

NO	Job Description	Work Schedule (Days)																		
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
I	Engineering drawing work																			
II	Material procurement work																			
A	Cutting marking production work																			
B	Machining production work																			
C	Feedup production work							Holiday							Holiday					
D	Welding production work																			
E	Finishing work																			
F	Quality control work																			
G	Painting work																			
H	Packing delivery work																			

2.1 Data Collection

Data collection is used to obtain data by observation and interviews[12]. Data collection in this study there were 2 respondents, namely production supervisor workers who participated in the project directly in the field. Data collection from the interview process indudes ducting project data implemented by PT. FTA and its owner from PT. MBI was planned to be completed in 14 days, while the project implementation was postponed to 19 days and identifying the work that was the cause of the project delay was interviewed with one respondent, namely the supervisor of the production section of the ducting project.

2.2 Data Processing

Data processing is carried out to analyze project delay factors and reschedule projects which include late work factors, identifying engineering drawing work and material procurement work using the Faut Tree Analysis (FTA) method for a combination of events that cause root problems that affect the project[13]. After a combination of basic event calculation of mocus analysis and minimum cut set, then reschedule the duration of project activities, analyze the calculation of early start, early finish, latest start, latest finish of each job and identify the critical path network diagram, determine the shortest rescheduling of work duration with alternatives to increase work hours and increase manpower, after that determine the relationship between the results of the fault tree analysis analysis with the results critical path method and compares with new scheduling so that optimal project time is obtained.

3. Results and Discussion

3.1. Fault Tree Analysis (FTA)

Data management is carried out using the Fault Tree Analysis (FTA) method which obtained the following steps.

3.1.1 FTA Steps Steps (Fault Tree Analysis)

Fault Tree Analysis consists of a combination of sequential faults that cause a failure event[14]. The first step is defining events in the system, creating an error tree from top event to intermediate event and top event logic gate relationships. The next step is to calculate the minimum cut set and analyze the FTA using the probability values of intermediate events to top events. Here can be seen FTA diagram figure 1.

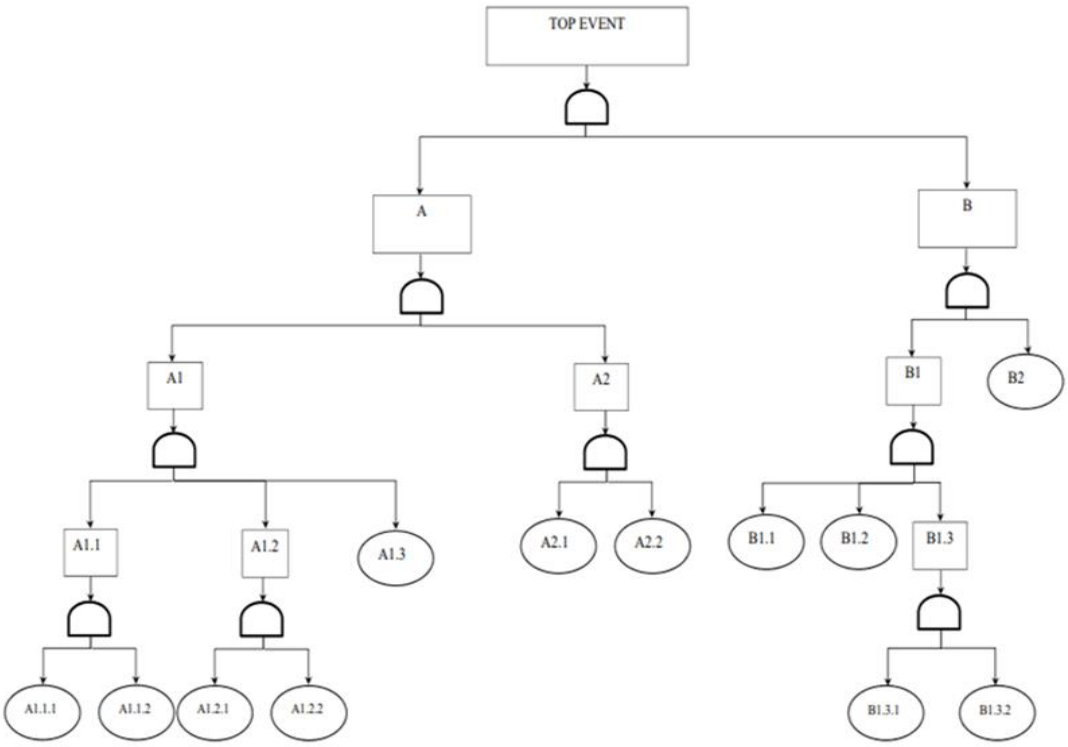


Figure 1. FTA Diagram Causes of Ducting Delays

3.1.2 Qualitative Analysis Fault Tree Analysis (FTA)

Analysis of the causes of delays in ducting projects at PT. FRA has two intermediate events from the interview results, namely drawing engineering work and material procurement work. In drawing engineering work, the problem is related to the owner's agreement, the main contractor and sub contractor cannot be agreed unilaterally when in the field, because they are still processing drawings with the owner's request to be reviewed again regarding their feedback. In material procurement work, it is still waiting for the owner's approval and drawing engineering work drawings, after the drawings come, the materials do not come immediately, because the company selects materials that propose prices in accordance with the project budget.

3.1.3 Quantitative Analysis Minimal Cut Set

Minimal cut set analysis is carried out with mocus analysis to find out the root of the problem in the fault tree[15]. Researchers distributed questionnaires to one respondent to find out the probability value and obtained 7 basic event mocus in drawing engineering work is A1.1.1; A1.1.2 ; A1.2.1 ; A1.2.2 ; A1.3 ; A2.1 ; A2.2 with a value of 0.4 x 0.4 x 0.3 x 0.3 x 0.3 x 0.3 x 0.2 = 0.0002592. In material procurement work, 5 basic events were obtained, namely B1.1; B1.2 ; B1.3.1; B1.3.2 ; B2 with a value of 0.4 x 0.3 x 0.3 x 0.4 x 0.4 = 0.00576. After conducting a Fault Tree Analysis (FTA) analysis, it is known that the event factors causing the delay in the ducting project and obtained 2 minimum cut sets, namely A = 0.0002592 and B = 0.00576 at PT. FRA can be seen in table 3.

Table 3. Event Fault Tree (Source PT. FRA)

Event	Information
A	Engineering drawing work
A1	Owner factor
A2	Maincont and subcont factors
A1.1	Slow owner in decision making
A1.2	lack of good coordination
A1.3	Owner's delay in reviewing feedback from maincont and subcont
A1.1.1	Waiting for the owner's decision regarding the drawing processing process

A1.1.2	Shop drawing release request
A1.2.1	Lack of coordination with maincont and subcont
A1.2.2	Late provision of drawing processing instructions
A2.1	Lack of coordination of subcont and maincont
A2.2	Re-drawing processing is carried out
B	Material procurement work
B1	Material selection factors
B2	Owner factor
B1.1	Propose a price according to the project budget
B1.2	Late PO fender list incoming material
B1.3	Delays in material delivery
B1.3.1	The materials ordered only came partially
B1.3.2	Not in accordance with the planned schedule

3.1.4 Analysis Sets a New Step

Researchers will determine a new step from table 3 event fault tree and a high probability value in engineering drawing work of 0,0002592 and material procurement work of 0,00576 to be combined which initially the factor causing the delay will be a new step factor, this way can optimize and will be usefull in the implementation of ducting projects in the company, then a new 4 steps will be obtained[16]. The first step is careful planning by creating a realistic project schedule with the project plan and identifying potential obstacle risks. The second step is to monitor project progress to track progress and changes to identify project delays early. The third step is effective communication between the company and the owner to provide periodic updates, monitoring work quality standards and the fourth step is planning schedule changes to anticipate if the planned project suddenly changes by communicating all parties in the project.

3.2. Critical Path Method (CPM)

The Critical Path Method can estimate the duration in the implementation of project activities that obtain supervision efficiently so that they are completed according to planning[17]. The critical trajectory of engineering drawing work activities with Early Finish (EF) = 2 days and Latest Finish (LF) = 4 days occurs slack 2 days. While EF material procurement work = 2 days and LF = 4 days there was a 2-day slack. Daily productivity calculation to compare production results with resources obtained value of volume of engineering drawing work = 4920.32 kg / 19 days = 258.964 kg / day.

3.2.1 Alternative Work Acceleration with Additional Working Hours (Overtime)

An alternative to speeding up project completion is the alternative approach of testing and implementing increasing overtime work hours[18]. Normal working hours on the project are 08.00-16.00 or 7 working hours and 1 hour rest at 12.00-13.00, additional ducting project working hours are 4 hours of work at 16.00-20.00. Calculation of additional productivity of drawing technique man-hours with daily productivity = 258.964 kg x 0.6 product reduction coefficient x 4 duration of overtime hours = 621.5136 kg. The calculation of the acceleration of adding working hours can be seen in table 4.

Table 4. Calculation of acceleration in addition to working hours

No	Job Description	Indeks (%)	Koef (-) Production	Overtime Hours	Prod. Planning
1	Engineering drawing work	258,964	0,6	4	621,5136
2	Material procurement work	170,328	0,6	4	408,7872
3	Cutting marking production work	134,901	0,6	4	323,7624
4	Machining production work	118,834	0,6	4	285,2016
5	Feedup production work	107,520	0,6	4	258,048
6	Welding production work	130,141	0,6	4	312,3384
7	Finishing work	104,332	0,6	4	250,3968
8	Quality control work	110,442	0,6	4	265,0608
9	Painting work	116,674	0,6	4	280,0176

10	Packing delivery work	144,400	0,6	4	346,56

Crash duration calculation is used to accelerate the project by maximizing resources and time on the project with volume = 4920.32 kg / productivity acceleration working hours = 621.5136 kg = 8 days crash duration.

3.2.2 Alternative Job Speed With the Increase of Manpower

The alternative of acceleration project completion by adding manprower to the ducting project is assumed to be 60% of the increase in daily productivity due to the addition of working hours = 621.5136 kg – Normal daily productivity = 258.964 kg / 258.964 kg x 100% = 60% or 0.6. Calculation of labor increase = 60% x 7 people = 4 people. Normal daily productivity calculation = 258,964 kg + $\left(\frac{258{,}964 \text{ kg x } 4 \text{ person}}{7 \text{ person}}\right)$ = 406,943 kg. Here is table 5 productivity acceleration labor addition.

Table 5. Accelerated Productivity After Labor Addition

No	Job Description	Indeks (%)	Early workforce	Labor Addition	Prod. Perc.
1	Engineering drawing work	258,964	7	4	406,943
2	Material procurement work	170,328	2	1	267,658
3	Cutting marking production work	134,901	5	3	211,987
4	Machining production work	118,834	2	1	186,739
5	Feedup production work	107,520	5	3	168,960
6	Welding production work	130,141	6	4	204,507
7	Finishing work	104,332	6	4	163,950
8	Quality control work	110,442	4	2	173,551
9	Painting work	116,674	3	2	183,344
10	Packing delivery work	144,400	3	2	226,914

The calculation of crash duration of work with alternative addition of labor is obtained from the calculation of volume = 4920.32 kg / 406.943 kg productivity acceleration of labor addition = 12 days crash duration.

3.2.3 Normal Cost and Crash Cost Alternative Increase in working hours and labor

Normal cost are the direct costs of completing each project work[19]. In normal times each project is assumed the normal labor cost per day is Rp. 100,000 x 7 initial labor x 2 days = Rp. 1,400,000. Crash cost of adding 4 hours of work for each job is assumed to be Rp. 150,000 + Rp. 100,000 normal daily cost = Rp. 250,000. Calculation of overtime crash cost Rp. 250,000 x 8 days overtime duration x 7 workers = Rp. 14,000,000. While the crash cost of adding labor with normal costs Rp. 100,000 x 4 people adding labor x 12 days crash duration = Rp. 4,800,000.

3.2.4 Relationship of Fault Tree Analysis Results with Critical Path Method

Based on the results of FTA (Fault Tree Analysis) analysis, the main factor causing the delay in ducting projects is in drawing engineering work whose problems related to the agreement of the owner, main contractor and sub contractor, image processing cannot be agreed unilaterally when in the field, because the main contractor and sub contractorare still processing images in accordance with the owner's request and there is a relationship with CPM (Critical Path Method) analysis, there are 10 critical jobs that depend on drawing engineering work will cause project execution time to increase because drawing engineering work is on a critical trajectory and the longest work is delayed in ducting projects.

4. Conclusion

Based on the results of ducting project research PT. FRA uses the FTA method, there is the highest probability value in the basic event based on the calculation of the minimum cut set on engineering

drawing work of 0.0002592 and obtained 4 new steps for the project contractors. Meanwhile, by using the CPM method, an alternative was obtained to accelerate the addition of 4 hours of work (overtime), a duration of 25.00% or 6 days faster and a normal duration of 19 days to 13 days faster and an additional cost of 91.96%, which is IDR Rp. 21,500,000. Meanwhile, the acceleration of the addition of alternative workers, the acceleration of the duration is 38.75% or 5 days faster than the normal duration of 19 days to 14 days and additional costs of 178.57%, namely Rp. 31,200,000[20].

From this study, it can be conduded that the alternative option of adding overtime work hours is more effectively used as an alternative to obtain accelerated duration and generate optimal costs compared to alternative labor additions and the additional costs incurred are also not too large, however it is improtant to conduct further research by analyzing the factors causing the first degree of delay so that the research is more specific and conduct research from the owner's point of view to find out more detailed information about the project cost data.

References

[1] M. Dany Fauzi and S. Salim Dahda, "Implementasi Metode House of Risk Pada Evaluasi Keterlambatan Proyek Cable Tray Support di PT. SSS," vol. VIII, no. 4, pp. 6890–6899, 2023.

[2] S. Wahyuningsih, I. I. P. Mulyatno, and I. S. J. Sisworo, "Analisa Faktor Penyebab Keterlambatan Proyek Reparasi dengan Metode Fault Tree Analysis (FTA) dan Penjadwalan Ulang dengan Critical Path Method (CPM) pada Kapal MT. Alice XXV di Galangan Semarang," *Tek. Perkapalan*, vol. 11, no. 1, pp. 14–22, 2023.

[3] Y. M. Fitriani, D. Andesta, and H. Hidayat, "Analisis Risiko Kerusakan Pada Mesin Las FCAW Dengan Pendekatan Metode Failure Mode And Effect Analysis (Studi Kasus : PT. Swadaya Graha)," *J. Serambi Eng.*, vol. 7, no. 4, 2022, doi: 10.32672/jse.v7i4.4663.

[4] F. D. Yanti and S. S. Dahda, "Penataan Layout Gudang Penyimpanan Material dan Peralatan Pendukung di PT . Swadaya Graha," vol. 20, no. 1, pp. 225–230, 2022.

[5] N. S. br Ginting and R. Kristiana, "Analisis Efektivitas Pengendalian Risiko Kecelakaan Kerja Pada Pekerja Proyek Konstruksi Menggunakan Metode Fine dan Fault Tree Analysis," *Teknik*, vol. 41, no. 2, pp. 192–200, 2020, doi: 10.14710/teknik.v39i3.20265.

[6] A. D. Pradiani and N. Nurhasanah, "Analisis Faktor Keterlambatan COD Proyek PLN UIP Kalbagtim dengan Metode Analytical Hierarchy Process dan Fault Tree Analysis," *J. Al-AZHAR Indones. SERI SAINS DAN Teknol.*, vol. 6, no. 1, p. 34, 2021, doi: 10.36722/sst.v6i1.639.

[7] H. D. Z. Putri, I. P. Mulyatno, and P. Manik, "Studi Manajamen Risiko dengan Metode FTA dan FMEA akibat Keterlambatan Proyek Pembangunan Kapal Perintis KM. Sabuk Nusantara 72," *Tek. Perkapalan*, vol. 8, no. 3, pp. 368–374, 2020.

[8] T. Iluk, A. Ridwan, and S. Winarto, "Penerapan Metode CPM Dan PERT Pada Gedung Parkir 3 Lantai Grand Panglima Polim Kediri," *J. Manaj. Teknol. Tek. Sipil*, vol. 3, no. 2, p. 162, 2020, doi: 10.30737/jurmateks.v3i2.1054.

[9] N. M. Astari, A. M. Subagyo, and K. Kusnadi, "Perencanaan Manajemen Proyek Dengan Metode Cpm (Critical Path Method) Dan Pert (Program Evaluation and Review Technique)," *Konstruksia*, vol. 13, no. 1, p. 164, 2022, doi: 10.24853/jk.13.1.164-180.

[10] B. Khrisdamara and D. Andesta, "Analisis Penyebab Kerusakan Head Truck-B44 Menggunakan Metode FMEA dan FTA (Studi Kasus : PT. Bima, Site Pelabuhan Berlian)," *J. Serambi Eng.*, vol. 7, no. 3, 2022, doi: 10.32672/jse.v7i3.4255.

[11] A. Lestari and N. A. Mahbubah, "Analisis Defect Proses Produksi Songkok Berbasis Metode FMEA Dan FTA di Home - Industri Songkok GSA Lamongan," *J. Serambi Eng.*, vol. 6, no. 3, 2021, doi: 10.32672/jse.v6i3.3254.

[12] A. Syarifudin and J. T. Putra, "Analisa Risiko Kegagalan Komponen pada Excavator Komatsu 150LC dengan Metode FTA DAN FMEA DI PT. XY," *J. InTent*, vol. 4, no. 2, pp. 1–10, 2021.

[13] R. Duyo, "Analisis Penyebab Gangguan Jaringan pada Distribusi Listrik Menggunakan Metode Fault Tree Analysis," *J. Tek. Elektro UNISMUH*, vol. 12, no. 2, pp. 1–12, 2020, [Online].

Available: https://journal.unismuh.ac.id/index.php/vertex/article/view/4017
[14] M. F. Perdana, B. Man, A. Hermawan, and H. S. Yudha, "Penerapan Keselamatan Kerja Dengan Metode Fault Tree Analysis Untuk Meningkatkan Produktivitas Kerja Di Pd . Sinar Mekar Application of Work Safety Using the Fault Tree Analysis (FTA) Method to Increase Work Productivity at UD . Blooming Light," vol. 13, no. 2, 2020.
[15] S. N. Trisaid, "Analisis Risiko Kecelakaan Kerja Pada Kegiatan Rig Service Menggunakan Metode Hirarc Dengan Pendekatan Fta," *J. Ilm. Tek. Ind.*, vol. 8, no. 1, pp. 25–33, 2020, doi: 10.24912/jitiuntar.v8i1.6343.
[16] A. T. Rahman, "Analysis of Inventory Control of Perishable Goods with Capital Constraints and Warehouse Capacity Using the Lagrange EOQ Method (Case Study UD. XYZ)," *Adv. Sustain. Sci. Eng. Technol.*, vol. 5, no. 3, p. 02303014, 2023, doi: 10.26877/asset.v5i3.17225.
[17] N. Sa'adah, E. Iqrammah, and T. Rijanto, "Evaluasi Proyek Pembangunan Gedung Stroke Center (Paviliun Flamboyan) Menggunakan Metode Critical Path Method (CPM) Dan Crashing," *Publ. Ris. Orientasi Tek. Sipil*, vol. 3, no. 2, pp. 55–62, 2022, doi: 10.26740/proteksi.v3n2.p55-62.
[18] W. Hermawan, "Single Sign On Using Keycloak Integrated Public Key Infrastructure for User Authentication In Indonesia's Electronic Based Government System," *Adv. Sustain. Sci. Eng. Technol.*, vol. 5, no. 2, p. 0230204, 2023, doi: 10.26877/asset.v5i2.15795.
[19] P. F. Dewi and Z. R. Kamandang, "Optimizing Project Performance by Applying the Crashing Method to Road Construction Project," *Adv. Sustain. Sci. Eng. Technol.*, vol. 5, no. 2, p. 0230203, 2023, doi: 10.26877/asset.v5i2.15944.
[20] S. H. Ali and D. Widyaningrum, "Quality Control Analysis With Statistical Quality Control (Sqc) and Failure Methodsmode Effect Analysis (Fmea) on the Production of Za Plus Fertilizer," *Adv. Sustain. Sci. Eng. Technol.*, vol. 6, no. 1, p. 0240104, 2023, doi: 10.26877/asset.v6i1.17358.

Advance Sustainable Science, Engineering and Technology (ASSET)
Vol. 6, No.2, April 2024, pp. 02402018-01 ~ 02402018-06
ISSN: 2715-4211 DOI: https://doi.org/10.26877/asset.v6i2.18701

Protein Concentrate From Tuna Head Waste Using Methanol-Acetone Solvent Extraction

Feri Ardiansyah*, Shofia Dwi Fitri Rahmasari, Kindriari Nurma Wahyusi

Departement of Chemical Engineering, Faculty of Engineering, Universitas Pembangunan Nasional "Veteran" Jawa Timur, Jln. Raya Rungkut Madya, Gunung Anyar, Surabaya, 60294, Indonesia

*feriardiansyahreal@gmail.com

Abstract. The frigate tuna (Euthynnus affinis) is known for its high protein content in the head, making it suitable for the production of protein concentrate used in animal feed. This research aims to investigate the influence of adding a methanol-acetone mixture solvent, the duration of the extraction process, and the protein content in the resulting concentrate using the Kjeldahl method. The protein concentrate is produced through the maceration extraction method, where 30 grams of the frigate tuna head sample is mixed with a methanol-acetone solvent in various ratios (1:9; 3:7; 5:5; 7:3; 9:1) at a temperature of 50 °C and a stirring speed of 500 rpm, with extraction times ranging from 2 to 6 hours. Subsequently, filtration is performed, and the precipitate is dried using an oven at 100 °C for 30 minutes. The dried sample is then subjected to protein content testing using the Kjeldahl method. Research results indicate that both protein content and extraction yield values increase with the duration of the extraction process, while the water content decreases. The optimal result in maceration extraction is achieved with the methanol-acetone mixture (9:1) treatment and a 6-hour extraction time, yielding a protein content of 89.15 %, water content of 5.57 %, and an extraction yield of 23.86 %.Protein concentrate can be used as animal feed to increase protein needs. Animals that are given sufficient protein will fatten and make the animal healthy

Keywords: acetone, extraction, frigate tuna head, methanol, protein concentrate

(Received 2024-03-21, Accepted 2024-03-29, Available Online by 2024-04-25)

1. Introduction

Fish in the ocean contain different amounts of protein. for example milkfish has around 20-24% protein, catfish has around 12% protein, anchovies contain 18.83% protein and pufferfish have around 21.40% protein [1],[2],[3],[4]. Frigate tuna (Euthynnus affinis) is a type of small, elongated tuna characterized by a scaleless body and a rigid dorsal fin texture. Frigate tuna is highly sought after by consumers for use as food due to its excellent high protein content. The waters in Lamongan Regency are listed as a dominant fishery resource, and frigate tuna is a commonly found species in these waters.

According to the statistical data from the East Java Provincial Fisheries and Marine Affairs Office (DKP), the production of frigate tuna increased from 47,083.85 tons in 2019 to 52,747.57 tons in 2020.

The significant increase in production is not accompanied by information on waste processing. The utilization of waste from frigate tuna heads is also suboptimal as only 50% of the frigate tuna meat is consumed, with the rest being discarded as waste [5]. The waste from frigate tuna heads has a protein content of 60.56%, making it suitable for use as protein concentrate. The processing of frigate tuna into protein concentrate is carried out through the maceration method with the addition of an organic solvent.

According to previous research conducted by Manullang in 2018, the production of fish flour from catfish head material containing 27% protein through a boiling process resulted in a protein content of 23.25% [6]. Another study by Apriyana in 2014 reported that the production of fish flour from catfish head material yielded a protein content of 9.97% [7]. Research conducted by Widiyanto in 2018, focusing on the production of fish flour from catfish head material, found a protein content of 19.66% [8].

The duration and use of solvents significantly influence the final outcome of the protein concentrate product. Protein concentrate is a powdered product with a protein concentration above 20%. It is produced through the removal of fat and oil components, resulting in a high protein content. The longer the extraction time, the more fat and oil are extracted, leading to an increase in protein content [9]. It is expected that the addition of a methanol-acetone mixture solvent during the extraction process will result in a higher quantity of extract over time. The produced extract has qualities and characteristics that can be observed based on water and protein content, thereby affecting the quality of the protein concentrate as animal feed [10].

The objective of this research is to examine the influence of adding a methanol-acetone mixture solvent and the duration of the extraction process from frigate tuna head waste. Additionally, it aims to assess the protein content in the protein concentrate obtained from frigate tuna head waste using the Kjeldahl method.

2. Methods

2.1. Research Method

The materials used are waste from frigate tuna heads obtained from Siwalankerto Market, Surabaya, and a methanol-acetone solvent obtained from a chemical supply store located on Tidar Street, Surabaya.

2.2. Procedures

In the production of modified protein concentrate (Siagian, 2019), waste from frigate tuna heads is washed, cut into small pieces, and dried in an oven for 120 minutes at a temperature of 100°C. It is then ground and sieved to a size of 40 mesh, and 30 grams are weighed [11]. Extraction is performed using the maceration method with a 1:3 ratio between the material and a methanol-acetone mixture solvent according to the specified variables (1:9; 3:7; 5:5; 7:3; and 9:1), totaling 90 ml, over variable time periods (2; 3; 4; 5; and 6 hours), with stirring at 500 rpm at a temperature of 50°C. The mixture is then filtered to separate the liquid and sediment. The sediment is dried in an oven at 100°C for 30 minutes. Subsequently, grinding and sieving are performed with a 40-mesh size. The resulting protein concentrate is then analyzed for protein content using the Kjeldahl method.

Different solvent ratios will get different yield results, the more polar the solution in the container will increase the yield obtained [12]. The longer the extraction time will increase the amount of yield obtained, this is because the contact time between the solvent and the material is longer [13]. The recommended temperature in the extraction process is 50°C, because too high a temperature will damage the material and will cause the solvent to evaporate more quickly. The best results are at an extraction temperature of 50°C [14].

Protein content analysis based on SNI No. 01-2354.4-2006 using the Kjeldahl method consists of three stages: destruction, distillation, and titration. In the modified protein content analysis [15].

(a) Destruction

100 mg of the sample is placed in a 100 ml Kjeldahl flask, added with 4.875 g potassium sulfate, 0.075 g copper sulfate, 0.05 g selenium, and 15 ml concentrated sulfuric acid. The mixture is shaken until homogeneous and heated in an acid cabinet until the fumes cease for approximately 2 hours. Heating is continued until boiling, and the clear liquid is maintained for 30 minutes and allowed to cool.

(b) Distillation

The destruction result is distilled by slowly adding 50 ml of 40% NaOH, then adding distilled water up to 100 ml. I tis then heated to boiling until the liquid is mixed. The distillation result is collected in an Erlenmeyer flask containing 15 ml of 4% boric acid and 3-5 drops of methyl red indikator.

(c)Titration

The distillation result is titrated with 0,1 N hydrochloric acid. The titration endpoint is marked and completed by a color change to pink. Titrations are performed three times to obtain a constant concentration. The formula for determining the protein content is :

$$\frac{(V2-V1)\times N\ HCl\ x\ BE\ N\times FK}{Sample\ weight\ \ (mg)}\times 100\% \qquad (1)$$

where:

V1 = Volume of blank HCl,

V2 = Volume of sample HCl,

N HCl= Normality of HCl,

BE N = Nitrogen equivalence (14.007),

FK = Conversion Factor (6.25).

3. Results and Discussion

The research material, which is waste from frigate tuna heads, underwent analysis to determine its protein content using the Kjeldahl method, water content using the gravimetric method, and Cd and Pb levels using the AAS (Atomic Absorption Spectrophotometry) method at the Laboratory of the Institute for Industrial Research and Consultation, Ketintang, Surabaya. The analysis results of the frigate tuna head waste are as follows:

Table 1. Results of the analysis of frigate tuna head waste.

Material	**Type of Analysis**			
	Protein content	Water content	Pb metal	Cd metal
Frigate tuna head flour	60,56%	17,80%	0,35 mm/kg	0,13 mm/kg

From Table 1, it is compared with the SNI 01-2715-1996 regarding Grade II fish meal with a minimum protein content of 55% and a maximum water content of 12%. According to SNI 2729:2013, the permissible consumption of heavy metals Cd and Pb in animal feed has a maximum limit of 0.5 mg/kg for Cd and 0.4 mg/kg for Pb. This complies with the Indonesian National Standard, indicating that frigate tuna head flour can be utilized for protein concentrate production.

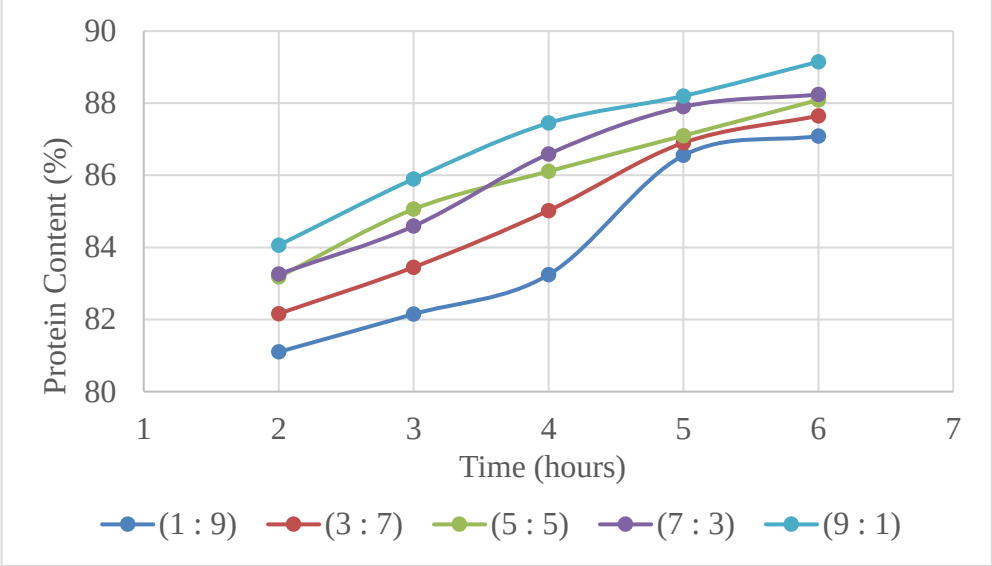


Figure 1. Relationship between Extraction Time and Protein Content

Based on Fig.1, the graph shows the relationship between extraction time and protein content, with protein content ranging from 81.1% to 89.15%. The highest protein content was achieved with the methanol-acetone mixture (9:1), reaching 89.15%. This aligns with SNI 3148.3 2009, which specifies a minimum protein content of 30% for layer chicken feed ingredients. The high protein content is attributed to the prolonged contact time between the material and the solvent, allowing compounds like fats and oils to dissolve in the solvent, while a significant amount of protein precipitates in the material. According to Heriansyah, the ability of each solvent to dissolve fats and oils is influenced by the polarity of the solvent. Solvents with high polarity (high dielectric constant) can dissolve more solutes. Methanol, with a dielectric constant of (33), is more effective in binding fat and oil compounds compared to acetone, which has a dielectric constant of (21) [16].

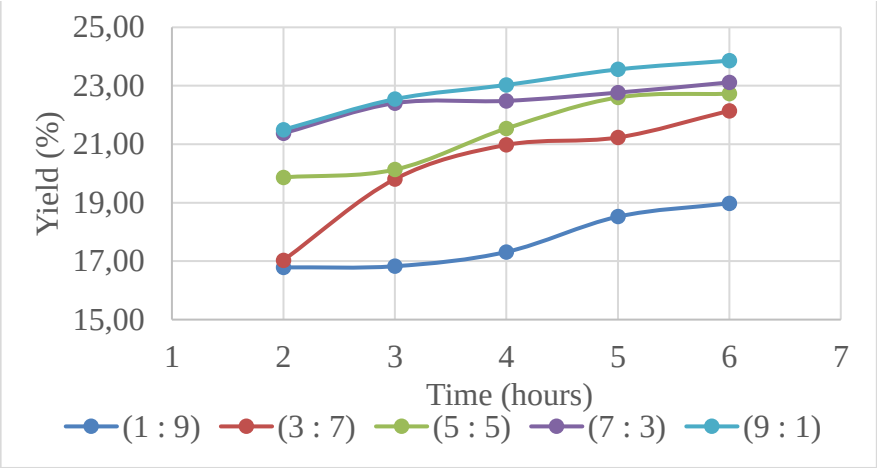


Figure 2. Relationship between Extraction Time and Yield

Based on Fig.2, the graph depicts the relationship between extraction time and yield, with the yield ranging from 16.79% to 23.86%. The highest yield was obtained with the methanol-acetone mixture (9:1), reaching 23.86%. A yield value is considered good if it exceeds 10%. The elevated yield value is attributed to the prolonged contact time, allowing more material to be extracted in the solvent. Methanol, having a higher boiling point (64°C) than acetone (56°C), does not evaporate quickly during the extraction process. This enables it to bind more fats and oils in the material compared to acetone. Stirring during extraction aims to generate a higher yield since the contact between the material and the

solvent occurs more frequently. According to Ningsih (2015), if the material is soaked for too long, more components such as fats that dissolve in the solvent will increase. The longer the extraction time, the higher the yield value obtained. Methanol, as one of the polar solvents, enhances the yield obtained because it can dissolve all organic compounds, contributing to the higher yield value [17].

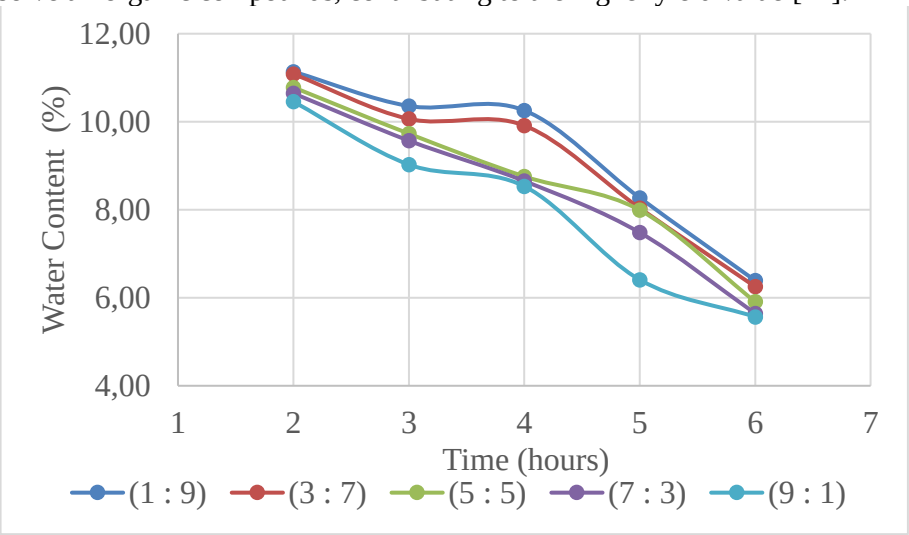


Figure 3. Relationship between Extraction Time and Water Content

Based on Fig.3, the graph shows the relationship between extraction time and water content, with the water content ranging from 5.57% to 11.14%. The optimal water content was achieved with the methanol-acetone mixture (9:1), which is 5.57%. This aligns with SNI 3148.3 2009, which specifies a maximum water content of 14% for layer chicken feed ingredients. The low water content is attributed to the prolonged extraction time, resulting in the evaporation of a significant amount of water vapor and causing a decrease in the material's water content. According to Manfaati (2019), the water content tends to decrease with the duration of the extraction process, This is because the pores in the material open, and the air density becomes loose, facilitating the evaporation of water from the material [18]. According liur (2020) suggests that water content is a determining factor in the shelf life of food materials. Higher water content in feed materials leads to high moisture, promoting bacterial growth and causing the feed material to spoil easily [19],[20].

4. Conclusion

Based on the conducted research, it can be concluded that the extraction time and the mixture of solvents influence the quality of the protein concentrate product. The best results were obtained with a ratio of 9:1 and an extraction time of 6 hours, yielding a protein content of 89.15%, water content of 5.57%, and yield of 23.86%. The protein concentrate in this research can be used as animal feed to fatten livestock. It is hoped that for future research, the heavy metal levels in tuna heads will get a value of 0 so that animal feed will be healthier.

Acknowledgements

Thanks are addressed to Mrs. Ir. Kindriari Nurma Wahyusi, MT. as a supervisor who has assisted in carrying out research both in the form of thinking and motivation. Hopefully this research can open up new and useful insights in various fields.

References

[1] M. Nurirmala. "Profil Protein Ikan Gabus (Channa striata), Toman (Channa micropeltes), dan Betutu (Oxyeleotris marmorata)" JPHPI. vol.23, pp.548-557. March 2020

[2] S.D. Rahayu. "Physicochemical Characteristics of Catfish (Clarias sp) Sausage with Addition of Moringa Leaf Flour (Moringa oleifera)," Advance Sustainable Science, Engineering and Technology (ASSET). vo.4, pp.1-12, may 2022
[3] R. Wahyudi. " Profil Protein pada Ikan Tenggiri dengan Variasi Penggaraman dan lama Penggaraman dengan Menggunakan Metode SDS-PAGE," jurnal sains dan teknologi. vol.2, pp.34-41,march 2019.
[4] M.F. Akhmadi. "Fortifikasi Kalsium dan Fosfor pada Crackers dengan Penambahan Tepung Tulang Ikan Bandeng (Chanos chanos)," Jurnal Ilmiah Perikanan dan Kelautan. vol.11, pp.49-54, april 2019
[5] Y. Sitompul, I. Sugitha and A. Duniaji. "Pengaruh Lama Perendaman Dalam Air Perasan Buah Belimbing Wuluh (Averrhoa bilimbi Linn) Dan Lama Penyimpanan Terhadap Karakteristik Ikan Tongkol (Euthynnus affinis) Pada Suhu Ruang," Jurnal Ilmu dan Teknologi Pangan, Vol. 9, pp.71-80, March 2020.
[6] Y. Manullang, L. Santoso and Tarsim, "Pengaruh Substitusi Tepung Ikan Dengan Tepung Kepala Ikan Patin (Pangasius sp) Terhadap Pertumbuhan Ikan Lele Sangkuriang (Clarias sp)," Jurnal Akuakultur Rawa Indonesia, Vol. 6, pp.129-140, December 2018.
[7] I. Apriyana, "Pengaruh Penambahan Tepung Kepala Ikan Lele (Clarias sp) Dalam Pembuatan Cilok Terhadap Kadar Protein dan Sifat Organoleptiknya," Unner Journal of Public Health. Vol. 3, pp.1-9, March 2014
[8] R. Widiyanto, R. Karnila, and M. Ilza, "Analisis Komposisi Kimia Tepung Kepala Ikan Lele Dumbo (Clarias gariepinus)," Jurnal Perikanan dan Kelautan. vol. 1, pp.1-8, March 2018
[9] Juwairiyah "Pembuatan Konsentrat Protein Dari Daging Bekicot Dengan Proses Ekstraksi" Jurnal Konsentrat Protein. vol. 2, pp. 35-41, December 2011
[10] A.A.D. Iffah, "Skrining metabolit sekunder pada sirip ekor Hiu Carcharhinus melanopterus," Jurnal kelautan dan perikanan, vol. 5, pp.65-72, april 2018
[11] R. M. Siagian,, M. Ilza and M. Sukmiwati, "Pembuatan konsentrat protein daging ikan patin (Pangasius hypophthalmus) menggunakan pelarut etanol dengan konsentrasi berbeda" jurnal perikanan dan kelautan. vol.1, pp.1-10, December 2019.
[12] Widyastutik, "Optimasi Perbandingan Pelarut dan Lama Maserasi terhadap Kadar Total Antosianin Ekstrak Jantung Pisang (Musa acuminata x Musa balbisiana)," Jurnal Farmasi Indonesia. vol.19, pp.167-175, April 2022.
[13] I.K.W. Putra, "Pengaruh Perbandingan Bahan dengan Pelarut dan Waktu Maserasi terhadap Ekstrak Kulit Biji Kakao (Theobroma cacao L.) sebagai Sumber Antioksidan," Jurnal Rekayasa dan Manajemen Agroindustri. vol.8, pp.167-176, Juni 2020.
[14] S. Chairunnisa, "Pengaruh Suhu dan Waktu Maserasi terhadap Karakteristik Ekstrak Daun Bidara (Ziziphus mauritiana L.) sebagai Sumber Saponin," Jurnal Rekayasa dan Manajemen Agroindustri. Vol.7, pp.551-560. March 2019.
[15] A. Nasution. "Penetapan Kadar Protein Pada Nanas Segar Dan Keripik Nanas Dengan Metode Spektrofotometri Uv-Vis Dan Kjehdahl," Journal Of Pharmacy and Science, vol. 4, pp.6-11, December 2020.
[16] I. Heriansyah, "Pengaruh jenis pelarut ekstraksi berbeda terhadap kandungan asam amino konsentrat protein teripang keling. Jurnal perikanan dan kelautan," vol. 1, pp.1-9, March 2018
[17] G. Ningsih, "Pengaruh Lamanya Waktu Ekstraksi Remaserasi Kulit Buah Durian terhadap Rendemen Saponin dan Aplikasinya sebagai Zat Aktif Anti Jamur," Jurnal konversi, vol. 4, pp.8-16, April 2015.
[18] R. Manfaati, H. Baskoro and M. Rifai, "Pengaruh Waktu dan Suhu Terhadap Proses Pengeringan Bawang Merah Menggunakan Tray Dryer," Jurnal Fluida, 12(2), pp.43-49, march 2019.
[19] L.J. Liur. "Kualitas Kimia dan Mikrobiologis Daging Ayam Broiler Pada Pasar Tradisional Kota Ambon," Journal of Biology and Applied Biology, vol. 3, pp.59-66, December 2020.
[20] A. Daud. "Kajian Penerapan Faktor yang Mempengaruhi Akurasi Penentuan Kadar Air Metode Thermogravimetri," Latjanus,v ol.24, pp.11-16, December 2020.

Advance Sustainable Science, Engineering and Technology (ASSET)
Vol. 6, No.2, April 2024, pp. 02402019-01 ~ 02402019-06
ISSN: 2715-4211 DOI: https://doi.org/10.26877/asset.v6i2.18623

MERN Implementation in Online Quiz Applications to Recognize and Avoid Social Media Hoaxes

Khairun Nisa Cipta Hapsari*, Yuri Yudhaswana Joefrie, Syaiful Hendra, Mohammad Yazdi Pusadan, Yusuf Anshori

Information Technology Department, Faculty of Engineering, Tadulako University, Jl. Soekarno Hatta No KM 9 Palu 94148, Central Sulawesi, Indonesia

*nisacipta1413@gmail.com

Abstract. Social media makes it easier for people to get information on the internet by allowing them to do more than just consume it. They can also create, comment on, and share content in different formats. False information that has permeated society can be readily ingested and used as a source of reference. A JavaScript application developed with the MERN stack technology, which combines several technologies—MongoDB, ExpressJS, ReactJS, and NodeJS—all functioning with one programming language. Data from numerous relevant research sources is gathered using qualitative approaches, and the Scrum method is employed to construct the application system. In addition to fostering creativity and innovation, this approach can resolve challenging and adaptive challenges. There is a maximum duration for every Scrum event. Once a sprint has begun, it cannot be stopped or shortened. Another event may be called off if its objectives have been fulfilled. The purpose of this study is to develop a website that offers pertinent tools and data to social media users in an attempt to stop the spread of hoaxes and fake news. The achieved result is that this application provides various quizzes designed to test users' knowledge of hoaxes and online scams.

Keywords: Scrum, Hoax, JavaScript, MERN, Quiz, Social Media

(Received 2024-03-21, Accepted 2024-03-29, Available Online by 2024-04-25)

1. Introduction

False information and fake news are widespread in society, causing much concern and loss. One example is the free internet news link from the government [1]. An example of a hoax circulating in society, such as free internet, caught the attention of 71.9 percent of respondents [1]. Furthermore, 36.7 percent of respondents said they would ask their friends or family about the information they received. However, there were 32.3% of respondents who did nothing about the information they received at that time [1]. Another example is the Indonesian Ministry of Communication and Information Technology finding 554 fake news about COVID-19 spreading widely until April 18, 2020 [2]. Another example of scams is email scams often promising prizes or attractive offers, leading people to be deceived and lose

money or sensitive personal data [3]

Social media allows users not only to consume content but also to participate, create, comment on, and disseminate content in various formats, making access to information on the internet easier now [4]. Hoaxes that have spread in society in the form of information can be easily consumed and used as reference material [5]. The impact on society can form a negative public opinion and create anxiety within the community [6]

The application created utilizes the MERN stack technology, which is a combination of various technologies working with one programming language, namely JavaScript [7]. MERN (MongoDB, ExpressJS, ReactJS, NodeJS), MongoDB is a document-oriented database, not a relational database. The reason for using a NoSQL database is to make it easier to scale in storing data. Document-oriented databases replace the concept of "rows" with a more flexible model, namely "documents" [8]. Express JS is a framework installed on top of the Node.js web server function to make APIs easier to use and add useful new features [9]. ReactJS is a front-end library developed by Facebook and used to support web frameworks. One of the advantages of ReactJS is allowing developers to create more interactive, stateful, and reusable user interface components. Other advantages include speed, ease, and scalability. In the Model View Control (MVC) paradigm, ReactJS is responsible for the view components [10] [11].

The aim of this research is to create a website that provides resources and relevant information for social media users with the aim of reducing the spread of fake news or hoaxes. Through interactive quizzes on this website, it is hoped that users can recognize and avoid the motives of hoaxes that can harm themselves or others.

2. Methods

2.1. Type of Research

Descriptive research is a type of research that aims to explain or describe a phenomenon or event in a specific and clear manner [12]. A qualitative approach is used to collect data from various related research sources. The process involves visual analysis and literature review to gain a deep understanding [13]. Data collection is carried out by conducting literature searches on journals, books, articles, news, and previous research related to the issue of hoaxes. The data obtained is then analyzed to serve as a source and basis for the preparation of quiz questions.

2.2. System Development

The application system development is carried out using the Scrum method. This method can solve complex and adaptive problems and generate creativity and innovation. The steps in the system development in the Scrum method include the stages of product backlog, sprint planning, sprint and daily scrum, and sprint review [14]. Each Scrum event always has a maximum duration. Sprints cannot be extended or shortened once started [15]. When the goals of another event have been achieved, another event can be terminated

3. Results and Discussion

3.1. Planning

In this stage, the main focus is to identify and plan the activities to be carried out within the sprint (a limited work period). Task allocation and prioritization of features to be developed are part of this process [17]. The application development team conducts planning, assesses user needs, and selects which features are the top priorities [18]. Each feature, such as the quiz system with levelling, educational articles, and the login and register landing page, will be broken down into smaller parts that can be implemented within a specific timeframe.

3.2. Analysis

In this stage, the writer identifies issues and challenges related to misinformation circulating on social media. This is done to gather the necessary data or information to serve as a foundation for research.

3.3. System Design

NoSQL is short for "Not Only SQL," which means that it doesn't solely use SQL syntax, such as MongoDB, which is based on documents. This means that it uses documents and collections instead of columns, rows, and tables. Since MongoDB is in JSON document form, data is stored in BSON (Binary JSON), which has data types like float, integer, string, date, boolean, etc. Due to this, MongoDB uses key-value pairs to store data, so each data has a key-value [19]. Because of its more flexible schema, MongoDB can automatically create its table structure when performing an Insert. MongoDB is a NoSQL database with larger data storage capacity and cheaper prices [20].

In Figure 1, the "user" collection is used to store data such as id, username, email, password, and level. This data is utilized when users attempt to log into the application. The "question" collection contains data about question id, question type, the question itself, answer options, answer key, and image links. This collection is employed to store all quiz questions from level one to level five. The "User answer" collection is utilized to store user id, question type, level, question id, user's answer, pass/fail status, and user's score. When users press the submit button, this data is stored in this collection. The relationship between the "user" collection and the "question" collection is marked by the presence of user_id and question_id. The "medal" collection is used to store level information, medals, and descriptions. The data in this collection provide information required when users successfully complete the quiz levels they are working on.

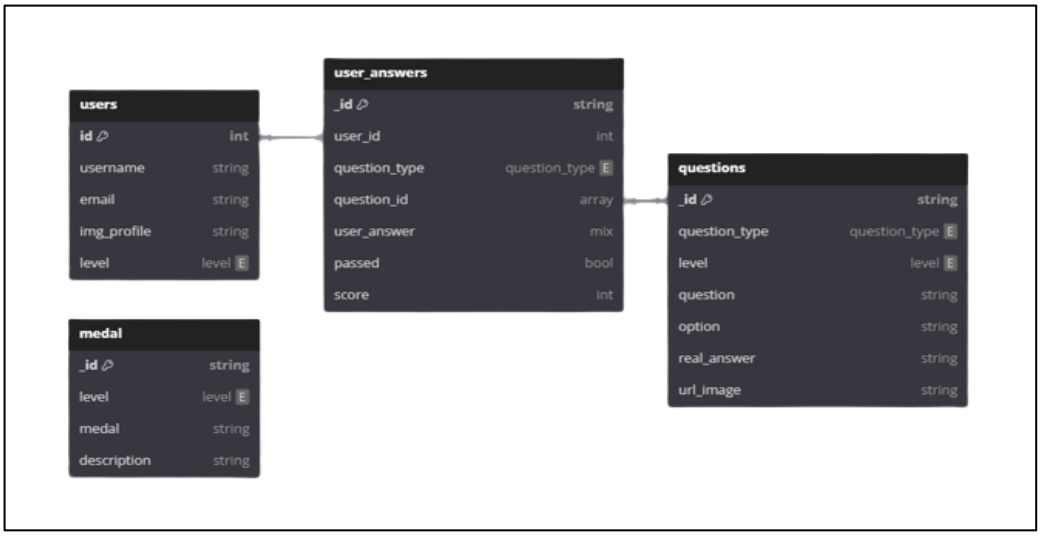


Figure 1. Relationship between collections

3.4. Implementation

The first step in introducing the application to users is the creation of a landing page. The landing page is designed to provide clear information about the application's purpose. This information includes a brief explanation of the benefits and utility of the application. To access the quiz, users are required to register first and then log in to the application. Additionally, the application presents relevant articles providing information and news about hoax-related topics and how to identify them. The implementation of MERN can be seen in Figure 2 below

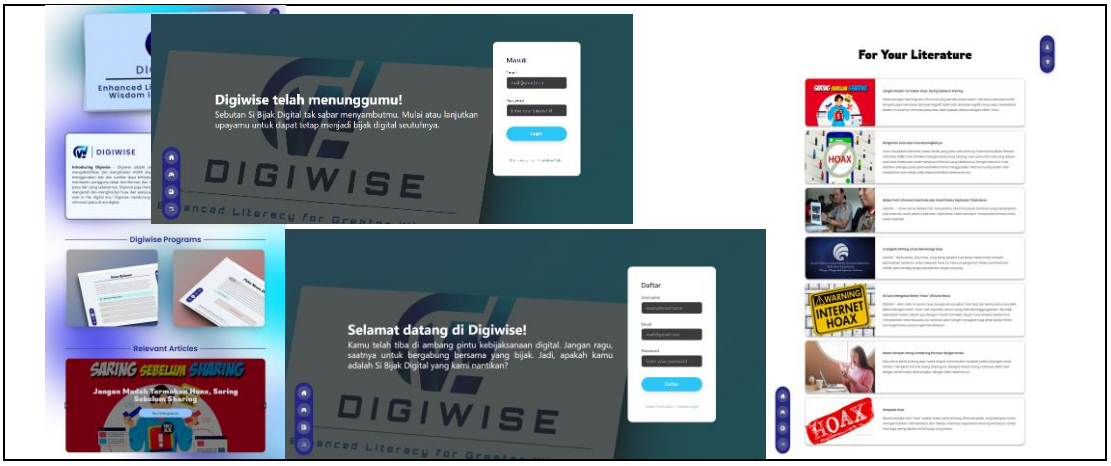


Figure 2. Landing page, User registration, User login, and Collection of articles

In this dashboard view, there are medals as visual indicators to provide users with instant information about their achievements in an easily understandable form. In Figure 3, there are visual differences in indicators in solving quiz questions in the application

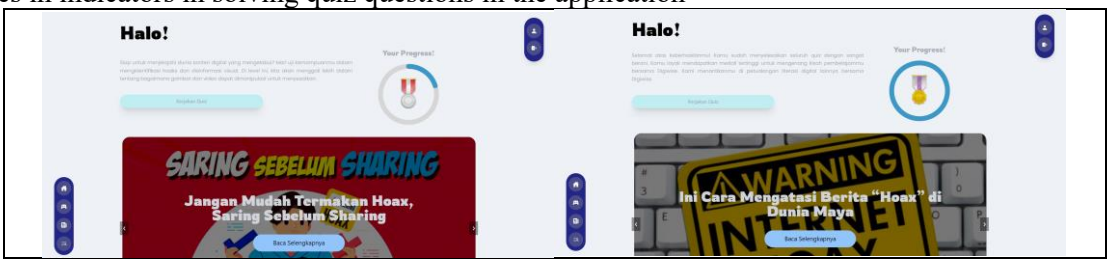
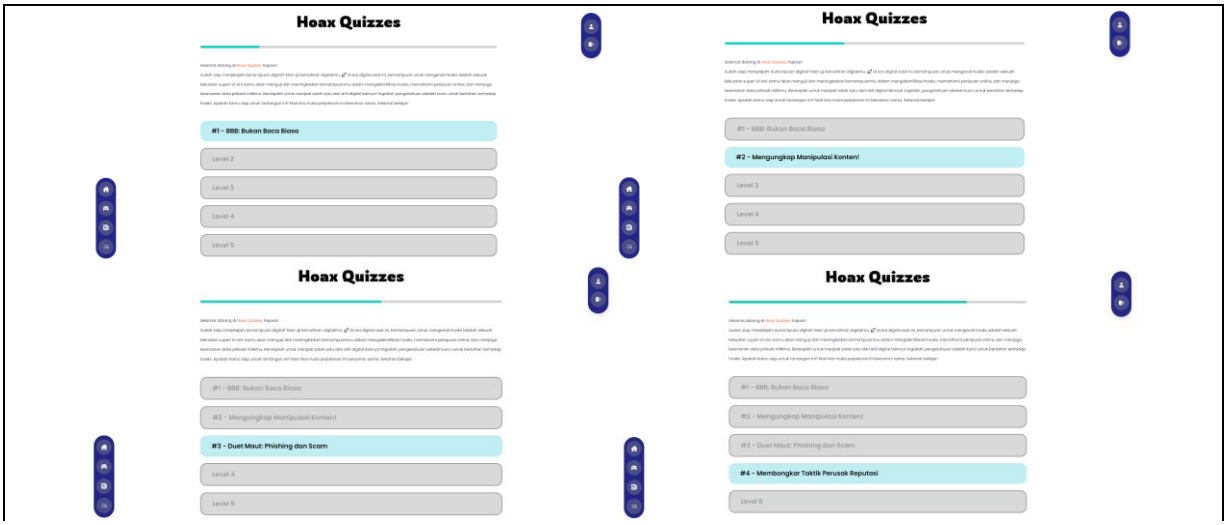


Figure 3. Quiz Dashboard Page

In Figure 4, it contains information about the user's progress level in the Digiwise application. On this page, users can view the list of levels that have been unlocked, showing their achievements in several stages. Additionally, this page also displays the levels that have not been unlocked yet.



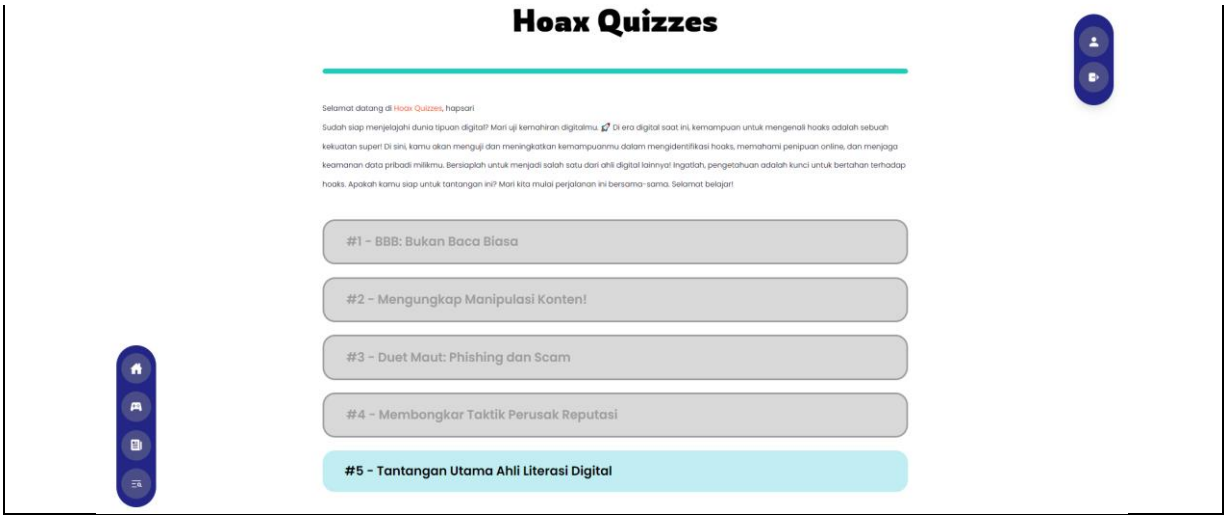


Figure 4. Level Information Page

This application provides various quizzes designed to test users' knowledge of hoaxes and online scams. For example, users can take multiple-choice quizzes that test their understanding of hoax characteristics and common scam tactics used on online platforms. For sample questions available in the application, refer to Figure 5 below

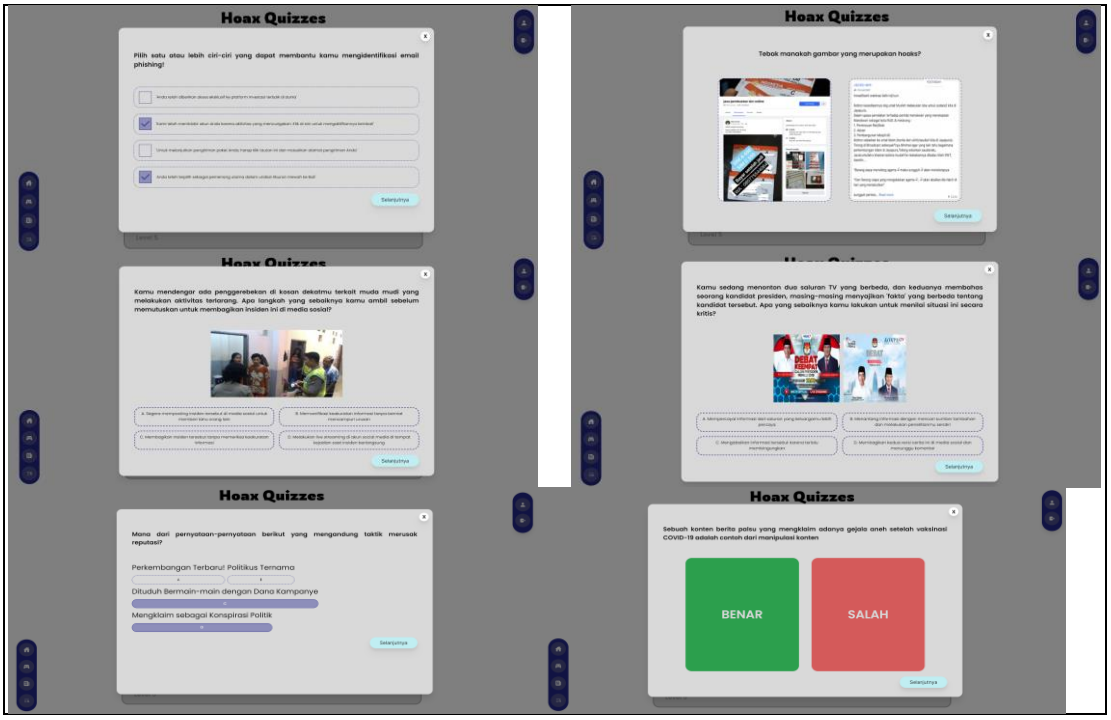


Figure 5. Overview of Questions at Each Level

Additionally, there are interactive quizzes that introduce users to real-life case studies of hoaxes that have spread on social media or specific websites. These quizzes aim not only to test users' knowledge but also to increase their awareness of the various strategies used by online fraudsters.

3.5. System testing

In this stage, the researcher will test the application to ensure its compliance with user needs in quiz execution. The goal is to identify bugs and ensure that the desired functionality runs smoothly.

Testing the API using Postman can check the API's response to each request, validate the returned data, and evaluate the overall system performance [21]. In this process, developers utilize Postman features to send HTTP requests to the API. Through Postman, developers can check the API's response. By leveraging Postman as an API testing tool, developers can ensure that the quiz answer checking process operates accurately and as expected in assessing user responses.

Table 1. System Testing Result

Test Case	Description	Result
Register	Register an account to log in	Success
Login	User login to access other features in the app	Success
Take the quiz	User can take existing quizzes	Success

In addition to testing through Postman, user feedback testing is also conducted using the SUS method. The System Usability Scale (SUS) method is a way to test the usability of an application. SUS was created as a "quick and dirty" usability measurement and is a questionnaire that can be used to measure how easily a computer system is used by users [22]. The SUS score interpretation scale can be seen in Figure 7.

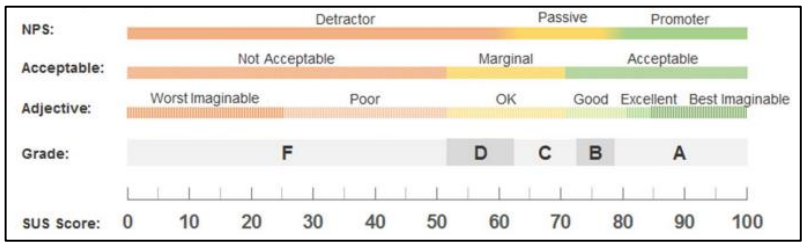


Figure 7. SUS Score Results Interpretation Scale

Responden	Pertanyaan												Jumlah	Nilai (Jumlah x 2,5)
	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	Q11	Q12		
R1	4	0	4	0	4	0	4	0	4	1	4	0	25	62,5
R2	3	0	3	1	3	0	4	0	3	1	4	0	22	55
R3	3	0	3	1	2	2	3	1	4	1	2	1	23	57,5
R4	3	1	3	1	1	1	1	3	1	2	2	1	20	50
R5	2	2	2	2	2	1	2	2	2	2	2	2	23	57,5
R6	3	1	3	2	2	1	2	1	4	2	2	2	25	62,5
R7	2	2	0	2	3	2	4	1	2	0	4	3	25	62,5
R8	3	2	2	2	2	2	2	2	2	1	2	2	24	60
R9	3	1	2	2	3	2	3	2	3	1	3	1	26	65
R10	4	0	3	0	4	0	4	0	4	0	4	0	23	57,5
R11	4	0	3	0	4	1	4	1	4	1	3	0	25	62,5
R12	3	3	2	2	2	3	3	1	3	2	2	2	28	70
R13	4	0	4	0	4	0	4	0	4	0	4	0	24	60
R14	4	0	2	1	3	1	4	0	4	1	4	0	24	60
Rata-Rata Score														60,18

Figure 8. SUS method test results

Based on figure 8, which contains 12 questions posed to users out of 14 respondents, the average score obtained after calculating the SUS method is 60.18. The average score falls into the "Ok" category and the "D" grade scale category when considering its alignment with the scores in figure 7. Based on the "adjective" value, obtaining an "OK" score implies that the application's impact in reducing the

spread of hoaxes has a positive effect. Despite its simplicity with a good user experience, this application can attract users to actively engage in the learning and understanding process of hoaxes, making them more vigilant and critical of the information they receive.

4. Conclusion

The main focus in this stage is identifying and planning activities for a sprint. The development team plans, considers user needs, and prioritizes features. Each feature is broken down into smaller components for implementation within a specific time frame. During this phase, the writer gathers data on misinformation challenges on social media. NoSQL, such as MongoDB, is used for data storage, employing key-value pairs and JSON-formatted documents. In the application, the "user" collection stores user data, "question" contains quiz questions, and "User answer" stores user submissions. The "medal" collection has level data, medals, and descriptions.

The application introduction begins with a landing page, providing information about the application's purpose and benefits. Users must register and log in to access the quiz. Relevant articles about hoax-related topics are also presented. The dashboard displays medals as visual indicators of user achievements. Figure 5 shows users' progress levels within the Digiwise application. The application offers various quizzes to test users' knowledge of hoaxes and online scams. Quiz questions cover hoax characteristics and common scam tactics. Interactive quizzes introduce users to real-life case studies of hoaxes.

There are shortcomings in the Digiwise application, which can certainly serve as directions for future research by potential improvements for the Digiwise application. This application lacks a feature for users to update their password when they forget it. Future research could consider the utilization of blockchain technology so that every published information will have a digitally verified cryptographic trace. Additionally, further development of the application could involve closer collaboration with institutions or organizations specializing in identifying and verifying the authenticity of information.

References

[1] A. Rizki Rangga Adi Negara & Indriani Rahmawati, “Status Literasi Digital di Indonesia 2021.” [Online]. Available: https://cdn1.katadata.co.id/media/microsites/litdik/Status_Literasi_Digital_diIndonesia%20_2021_190122.pdf

[2] V. Juliswara, F. Muryanto, P. Sosiologi, and S. Kartika Bangsa Yogyakarta, “Model Penanggulangan Hoax Mengenai Berita Covid 19 untuk Pengembangan Literasi Digital Masyarakat di Indonesia,” 2022. [Online]. Available: http://Jiip.stkipyapisdompu.ac.id

[3] T. M. Ajayi, “Discursive-manipulative strategies in scam emails and SMS: The Nigerian perspective,” vol. 18, no. 1, pp. 175–195, 2022, doi: doi:10.1515/lpp-2022-0008.

[4] P. Arsi, B.A. Kusuma, and A. Nurhakim, “Analisis Sentimen Pindah Ibu Kota Berbasis Naive Bayes Classifier,” *JURNAL INFORMATIKA UPGRIS* , vol. Vol 7, 2021.

[5] N. Aksin, S. Ag, M. Si, and S. Baedowi, “Berita Bohong (Hoax) Perspektif Hukum Islam,” 2020.

[6] N. Zaimatus Septiana and M. Wahyu, “Journal of Dedication Based on Local Wisdom DAMPAK BERITA HOAX PADA MASYARAKAT: STUDI FENOMENOLOGI KELURAHAN NGRONGGO KOTA KEDIRI,” vol. 1, pp. 207–216.

[7] N. H. Kamal and A. Gunaryati, “Implementasi Pengembangan Web Menggunakan Teknologi MERN Stack pada Sistem Informasi Akademik Siswa Berbasis Web,” vol. 11, no. 3, 2023, doi: 10.26418/justin.v11i3.

[8] S. Bradshaw, E. Brazil, and K. Chodorow, *MongoDB: the definitive guide: powerful and scalable data storage*. O’Reilly Media, 2019.

[9] R. Handoyo, L. W. Santoso, and A. Setiawan, “Real-Time BPMN Website Menggunakan Teknologi MERN Stack,” 2019. [Online]. Available: https://www.mongodb.com/blog/post/the-modern-
[10] J. Panjaitan and A. F. Pakpahan, “Perancangan Sistem E-Reporting Menggunakan ReactJS dan Firebase,” *Jurnal Teknik Informatika dan Sistem Informasi*, vol. 7, no. 1, Apr. 2021, doi: 10.28932/jutisi.v7i1.3098.
[11] M. de Sousa and A. Gonçalves, “humanportal – A React.js case study,” in *2020 15th Iberian Conference on Information Systems and Technologies (CISTI)*, 2020, pp. 1–6. doi: 10.23919/CISTI49556.2020.9141070.
[12] M. Jindan, H. R. Ngemba, S. Hendra, R. Laila, and S. Syahrullah, “Integrated Web-based Palu City Blood Donor Service Application Model Using ReactJS and ExpressJS,” *Advance Sustainable Science Engineering and Technology*, vol. 5, no. 3, p. 02303013, Oct. 2023, doi: 10.26877/asset.v5i3.17196.
[13] S. Hanyfah, G. Ryan Fernandes, I. Budiarso, and J. RayaiTengah Nomor, “PENERAPAN METODE KUALITATIF DESKRIPTIF UNTUK APLIKASI PENGOLAHAN DATA PELANGGAN PADA CAR WASH,” 2022.
[14] W. A. Prabowo and C. Wiguna, “Sistem Informasi UMKM Bengkel Berbasis Web Menggunakan Metode SCRUM,” *JURNAL MEDIA INFORMATIKA BUDIDARMA*, vol. 5, no. 1, p. 149, Jan. 2021, doi: 10.30865/mib.v5i1.2604.
[15] T. S. Ardan, “Perancangan Sistem Informasi Manajemen Aset Berbasis Web Menggunakan Metode Scrumban (Studi Kasus: Universitas Subang),” pp. 1–5, 2024.
[16] A. Izhar, “PENGEMBANGAN REST API DENGAN MENGGUNAKAN EXPRESS JS UNTUK MENCARI MENTOR PRIBADI,” *Jurnal Informatika Terpadu*, vol. 9, no. 2, pp. 92–102, 2023, [Online]. Available: https://journal.nurulfikri.ac.id/index.php/JIT
[17] W. Warkim, M. H. Muslim, F. Harvianto, and S. Utama, “Penerapan Metode SCRUM dalam Pengembangan Sistem Informasi Layanan Kawasan,” *Jurnal Teknik Informatika dan Sistem Informasi*, vol. 6, no. 2, 2020.
[18] Y. H. Wibowo and G. P. Mahardhika, “Migrasi Aplikasi Travel Aja React Native ke Kotlin Native Menggunakan Pendekatan Metode Scrum,” 2023.
[19] Yulius and Yeremia Alfa Susetyo, “View of Analisis dan Penerapan Database Mongodb pada Aplikasi Manajemen Dokumen di PT. XYZ,” 2023.
[20] B. Cahyo Santoso, Y. Natasya, S. Willian, and F. Alfando, “Tinjauan Pustaka Sistematis terhadap Basis Data MongoDB,” 2020. Accessed: Mar. 18, 2024. [Online]. Available: https://pdfs.semanticscholar.org/5acf/67b2b10f85a8d69c0ab84911a19c8839b345.pdf
[21] L. Dwinur Andrianto and D. Fatrianto Suyatno, “Analisis Performa Load Testing Antara Mysql Dan Nosql Mongodb Pada RestAPI Nodejs Menggunakan Postman,” 2024.
[22] Z. Miftah and I. P. Sari, “ANALISIS SISTEM PEMBELAJARAN DARING MENGGUNAKAN METODE SUS,” *Research and Development Journal of Education*, vol. 1, no. 1, p. 40, Oct. 2020, doi: 10.30998/rdje.v1i1.7076.

Advance Sustainable Science, Engineering and Technology (ASSET)
Vol. 6, No.2, April 2024, pp. 02402020-01 ~ 02402020-09
ISSN: 2715-4211 DOI: https://doi.org/10.26877/asset.v6i2.18502

Analysis of Improving the Quality of Fatty Acid Methyl Ester (FAME) Products Against Acid Value (AV) Levels Using the Six Sigma and Kaizen Methods

Mohammad Salva Afifi*, Said Salim Dahda

Faculty of Engineering, Universitas Muhammadiyah Gresik. Jl. Sumatra 11 GKB, Gresik 61121, Indonesia

*vavagresik82@gmail.com

Abstract. PT XYZ is a palm oil derivative processing company that produces biodiesel / FAME (Fatty Acid Methyl Ester) products. However, there are several obstacles that still occur during the company's production process which causes the acid value (AV) level of FAME products to not meet specifications and there are 3% defective products. Therefore, this study aims to determine the improvement plan by applying kaizen to the Six Sigma improvement stage and reducing outspec products in the FAME production process. The results of the research conducted using the six sigma method show that the company has a DPMO value of 71,459 and a sigma value of 2.96 which is in the average position of the industry in Indonesia based on the results of data processing. While recommendations for improvement by implementing kaizen include training machine operators, making standard operating procedures (SOPs) for machine inspections on each shift, reducing the amount of strearin added to AV raw materials before pre-treatment, maintenance and machine calibration. This step are expected to overcome the main problems identified in the study.

Keywords: Quality, Quality Control, Biodiesel, Six Sigma, DMAIC, Kaizen

(Received 2024-03-23, Accepted 2024-04-11, Available Online by 2024-04-25)

1. Introduction

Biodiesel, which is made from vegetable oils such as palm oil, castor oil, soybean oil, etc., is produced through a transesterification process with methanol and ethanol reagents, which is an environmentally friendly way to reduce the use of fossil fuels [1]. SNI 7182:2015 is the national standard for biodiesel quality in Indonesia. This standard requires an acid value (AV) of 0.4%, a water/moisture concentration of 340 ppm, a monoglyceride concentration of 0.525%, a methyl ester concentration of 96.5%, and a contaminant/impurity concentration of 20 ppm [2]. So, good quality comes from good processes, in accordance with predetermined quality standards based on market needs [3].

PT. XYZ is a palm oil derivative processing company that produces biodiesel/FAME (Fatty Acid Methyl Ester) products. PT. XYZ has FAME quality standards that have been set by the company and customers. However, there are several problems that still occur during the company's production

process, which causes the acid value (AV) levels of FAME products to not comply with specifications and there are 3% of products that are defective. Product quality is very important to consider because it greatly affects customer satisfaction with the products produced. To face this challenge, manufacturers are now required to be able to produce products that meet and even exceed customer expectations [4]. Overcoming these problems by analyzing quality control using the kaizen and six sigma methods [5].

Where the Six Sigma Method is a method that finds the source of problems in production and reduces the level of product defects to reduce costs and streamline production time [6]. By going through the DMAIC stage approach (Define, Measure, Analyze, Improve and Control) [7]. Then, the kaizen method is used to find the right improvement recommendations for each work component so that the improvement goals achieved can be sustainable [8]. The Kaizen approach looks at the business from a broader perspective. The goal is to improve every aspect of the business, setting process standards, increasing efficiency, and constantly looking for ways to reduce waste. It is a system that is always focused on improvement, from the entry level to the highest [9]. Kaizen, meaning continuous improvement, is one of the important ways to gain an edge in production and is essential in today's competitive environment [10]. Kaizen is a principle of continuous improvement represented by the Japanese terms KAI and ZEN, which mean "change" and "getting better." [11]. According to [12] four tools can be used to implement kaizen, namely:

1. Kaizen checklist
   Used to find problems and point out opportunities and improvements.
2. The five-step Kaizen strategy
   The 5-S movement, which is another Japanese word that starts with the letter S, namely Seiri (Sorting), Seiton (Organizing), Seiso (Cleaning), Seiketsu (Stabilizing), and Shitsuke (Habituation).
3. Five W and One H:
   Who, What, Where, When, Why, and How are tools that are widely used as management tools in a variety of environments.
   environments
4. Five M's checklist
   In the kaizen approach, there are five main components involved in every process: people, machines, materials, techniques, and measurements. Within each process, any improvements can be made by reviewing the parts of the production process.

2. Methods

This research uses data such as laboratory testing results and FAME product defect data to reduce the value of defective products and provide improvement recommendations for PT. XYZ. DMAIC and Kaizen methods are used. Researchers collected data samples from 01 October to 01 November 2023. Then this data was processed and analyzed using various DMAIC tools, including X and R maps, process capability, and fishbone diagrams[13]. In addition, the kaizen implementation tool M Checklist was also used [14]. Then the results are processed using the following stages:

A. Define (identify the problem)
   At the Definition stage, observations and interviews with the head of production are carried out to determine Critical to Quality (CTQ). Which is a quality characteristic of the product produced and to determine the specification limit of a product in accordance with customer desires [15].
B. Measure (measurement)
   In this step, process capability analysis is carried out by calculating the Defect Per Million Opportunities (DPMO) value to get the results of the sigma value calculation [16].
C. Analyze (analysis)
   In the analysis stage, a fishbone diagram is used to analyze the causes of the increase in FAME Acid Value (AV) levels. To do this, the fishbone diagram is divided into human, material, machine and method elements [17]. This analysis stage uses the findings from the interview with the production manager to produce the fishbone diagram results. The results of the analysis

stage will be used to make recommendations for improvement at the improvement stage.

D. Improve (repair)
After identifying the root cause of the quality problem, the fourth step of the Six Sigma approach is improvement. The improvement is by implementing kaizen using kaizen method M checklist [18]

E. Control
The final analysis process of Six Sigma is control, which emphasizes the documentation and dissemination of the actions that have been taken [19]

3. Results and Discussion

Data collection

The following is the data of FAME production defect products on 01 October – 01 November 2023

Table 1. FAME production defect products on 01 October – 01 November 2023

No	Date	Production (Ton)	Defect type Acid Value (Ton)	Defect products (%)
1	01 October – 01 November 2023	48.000	1.440	3

Table 1 shows that the total production is 48,000 tons, with defective products of 1,440 tons or about 3%.

The following are the results of quality measurements of Fatty Acid Methyl Ester (FAME) AV (Acid Value) levels carried out in the PT. XYZ laboratory on October 1 to November 1, 2023 used for this study. shown in table 2.

Table 2. Test Results for FAME Acid Value (AV) levels

Date	Acid Value (AV) levels		
01-Okt-23	0,129	0,191	0,187
02-Okt-23	0,217	0,196	0,125
03-Okt-23	0,178	0,155	0,151
04-Okt-23	0,199	0,106	0,204
05-Okt-23	0,222	0,188	0,135
06-Okt-23	0,143	0,161	0,130
07-Okt-23	0,201	0,210	0,231
08-Okt-23	0,198	0,259	0,210
09-Okt-23	0,177	0,213	0,189
10-Okt-23	0,180	0,143	0,124
11-Okt-23	0,164	0,222	0,207
12-Okt-23	0,200	0,165	0,156
13-Okt-23	0,142	0,152	0,163
14-Okt-23	0,280	0,258	0,242
15-Okt-23	0,210	0,143	0,160
16-Okt-23	0,132	0,123	0,142
17-Okt-23	0,211	0,221	0,190
18-Okt-23	0,132	0,146	0,190
19-Okt-23	0,178	0,193	0,278
20-Okt-23	0,267	0,221	0,175
21-Okt-23	0,259	0,211	0,193
22-Okt-23	0,293	0,268	0,220
23-Okt-23	0,190	0,156	0,172
24-Okt-23	0,190	0,134	0,145
25-Okt-23	0,170	0,145	0,153
26-Okt-23	0,123	0,134	0,164
27-Okt-23	0,135	0,152	0,190
28-Okt-23	0,298	0,279	0,200
29-Okt-23	0,135	0,123	0,150
30-Okt-23	0,150	0,142	0,167
31-Okt-23	0,153	0,153	0,124
01-Nov-23	0,163	0,191	0,145

Table 2 shows the results of the FAME Acid Value (AV) test conducted by the laboratory of PT XYZ.

Data processing

A. Define stages

At this stage, the first step taken is to determine the project to be implemented based on a predetermined priority scale. After that, the next step is to establish an action plan. After that determine Critical to Quality (CTQ). In this research, the factor or driver that consumers want is the level of AV (Acid Value) which produces product quality standards in the table 3.

Table 3. CTQ (Critical To Quality) FAME Products

CTQ (Critical to Quality)	Type of Content	Specification	Description
CTQ -1	Acid Value (AV)	AV <0,25%	One of the main factors influencing FAME quality is acid content. High acid levels can cause the resulting biodiesel to become corrosive or damage fuel injectors, damage fuel pumps and diesel engines.

Table 3 shows that the company set a CTQ for the Acid Value (AV) level of FAME products at <0.25%.

B. Measure Stages

At this stage, the Defect Per Million Opportunities (DPMO) value is calculated to obtain the sigma value calculation results. At PT. XYZ, this research will evaluate the performance of the Fatty Acid Methyl Ester (FAME) production process using Maps $\bar{X}$ and R, then value the process capability for each quality characteristic or Critical to Quality (CTQ).

1. Maps $\bar{X}$ and R

   Maps $\bar{X}$ and R control charts are used to illustrate the differences that occur in the production process [20]. Next, data processing is carried out using Minitab to calculate the CL center limit, UCL lower limit, and LCL lower limit values which are displayed in the graph. The following image 1 shows a maps of $\bar{X}$ and R Acid Concentration (AV)/Amount of Acid.

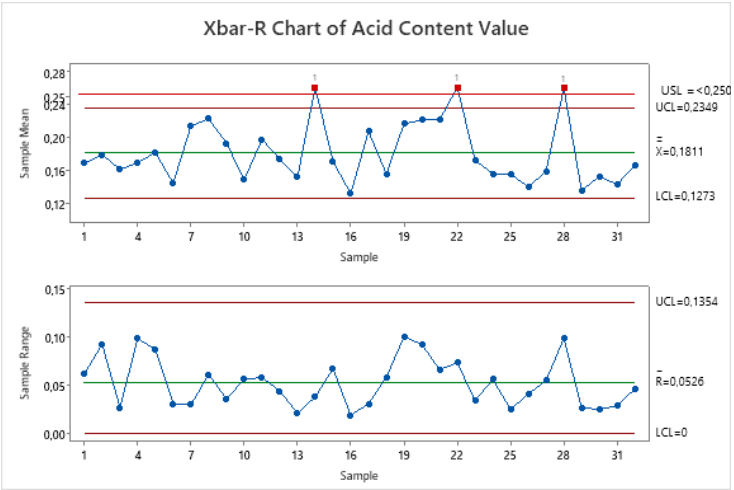


Figure 1. Map of $\bar{X}$ and R AV levels

As shown in Figure 1. On the $\bar{X}$ map, the CL value is 0.181, the UCL value is 0.235, and the LCL value is 0.127, data numbers 14, 22, and 28 on the $\bar{X}$ map are (outside the control limits). while on the R map, the CL value is 0.052, the UCL value is 0.135, and the LCL value is 0. Thus, all data on the R map are within (inside the control limits).

2. Processability Acid Content Value (AV)
   To improve process quality, the process capability index is used to measure the performance relationship between the actual process and the expected specification limits..

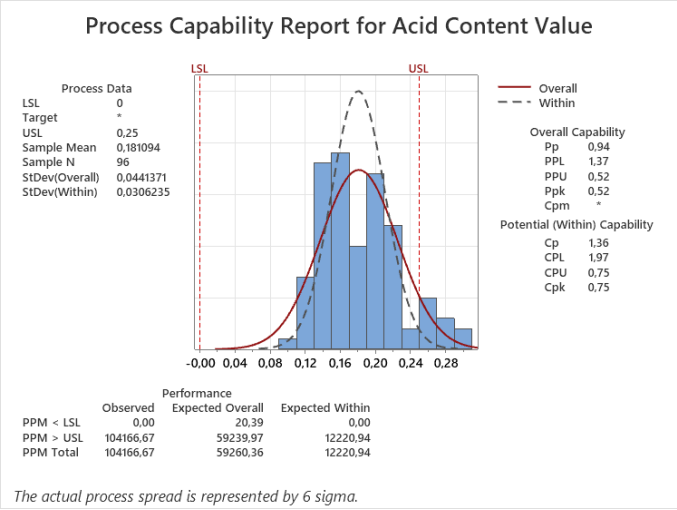


Figure 2. Process Capability AV Level

As shown in Figure 2. The results show Cp of 0,60, CPL of 0,90, CPU of 0.29, and Cpk of 0.29. Since this study is a short-term study, the goal is to find out the actual capability of the production process based on the quality of the products produced. If the Cpk value is <1, then the production process may fail to produce products that do not conform to specifications [1].

3. Sigma Value Measurement

   a. DPMO value
      - Part Per Million Defect (PPM) Overall Performance

Total DPMO
= PPM Overall + PPM Within (1)
= 59.239 + 12.220
= 71.459

The DPMO value is 71.459, and when converted to the Sigma Table, the sigma value is 2,96. The company is at the industry average position in Indonesia with this sigma value.

C. Analyze stages

Analysis of causal factors affecting low FAME quality levels using fishbone diagram. The following fishbone diagram shows the variation of acid content value (AV).

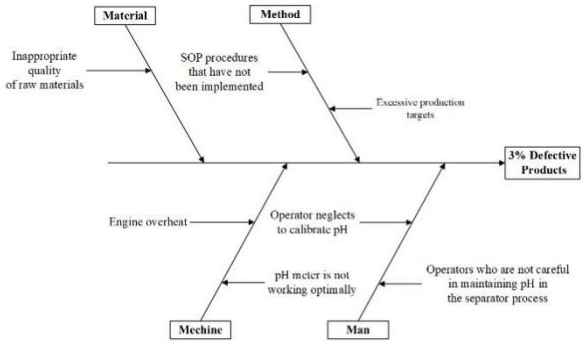


Figure 3. Fishbone Diagram of AV Levels

Based on Figure 3. The varying AV levels are caused by 4 (four) factors. The following is an explanation of the fishbone diagram of Acid Value (AV) levels :

Table 4. Explanation of AV Levels Fishbone Diagram

Category	Problems	Description
Man	Operators who are not careful in maintaining pH in the separator process	pH is one of the processes that can cause high AV levels in FAME. pH that is not in accordance with procedures can occur because the operator does not monitor properly.
	Operator negligence in Performing pH calibration	Due to fatigue, the operator fails to calibrate the pH meter. can cause the pH meter value does not match the pH of the field.
Method	SOP procedures that have not been implemented	The absence of standard operating procedures (SOPs) led to incomplete machine inspections, which resulted in some elements not being evaluated.
	Targeted production excess	One of the factors affecting the quality of FAME produced is excessive production targets. Excessive production targets cause the production process to be accelerated than usual, resulting in a decrease in FAME quality.
Material	Inappropriate quality of raw materials	The main factor affecting the quality of the FAME produced is the raw material. mixing RBDPO with Stearin this causes the AV value of the raw material to not match.
Machine	Engine overheat	Engine overheat occurs when the engine is used continuously. This reduces the performance of the engine significantly and may even cause the engine to become unusable.
	pH meter is not working optimally	The pH meter is not working properly because there is no pH inspection or calibration.

Table 4 shows the explanation of the fishbone diagram of AV levels caused by four defect-causing factors: man, machine, material, and method.

D. Improve Stages

At the improvement stage, the kaizen implementation tool used is the Five-M Checklist

Table 5. Five M checklists

Factor	Problems	Description
Man	- Operators who are not careful in maintaining pH in the separator process - Operator negligence in Performing pH calibration	- Strictly and gradually supervise the separator process - Provide training to separator machine operators on machine mechanisms and process procedures - Appoint an operator to be responsible for calibration in each shift
Method	- SOP procedures that have not been implemented - Targeted production excess	- Thorough machine inspections and regular scheduling of each shift - Reducing the production target so that the FAME production produced is in accordance with specifications
Material	- Inappropriate quality of raw materials	- Reduce the amount of strearin mixed so that the AV before the pre-treatment process meets the standard.
Machine	- Engine overheat - pH meter is not working optimally	- Perform routine maintenance of separator machines - Perform pH meter calibration every shift

Table 5 above shows suggestions that can help reduce production failures in FAME products by applying the Kaizen method

E. Control Stage

This control stage emphasizes the documentation and dissemination of actions that have been proposed for the improvement process :

a. Provide training to separator machine operators on machine mechanisms and process procedures
b. Making SOPs regarding routine machine inspections on each shift and Reducing the production target so that the FAME production produced is in accordance with specifications
c. Reduce the amount of strearin mixed in the raw materials so that the AV before the pre-treatment process meets the standard.
d. Perform routine maintenance of the separator machine and calibrate the pH meter every shift.

In addition, the company should have a strong internal audit team to ensure continuous improvement and up-to-date update documents for long-term success.

4. Conclusion

According to the above research, there are several factors that affect the quality of FAME products that do not meet specifications, such as human resource management, raw material handling, and production machine maintenance, which cause acid levels to increase during the production process. Based on data analysis and processing using the six sigma method, the company has a DPMO value of 71.459 and a sigma value of 2.96, which indicates that the company is in the average position of the

industry in Indonesia. Furthermore, suggestions for improvement are based on kaizen implementation tools : include training machine operators, making standard operating procedures (SOPs) for machine inspections on each shift, reducing the amount of strearin added to AV raw materials before pre-treatment, maintenance and machine calibration. However, for validation and deeper understanding, further research is needed with a wider dataset and longer research period.

References

[1] Ridho Dwi Syahputra, "ANALISIS SIX SIGMA DALAM PENINGKATAN KUALITAS PRODUK CRUDE PALM OIL (CPO) GUNA MEMINIMALISIR DEFECT," 2023.

[2] M. Busyairi *et al.*, "Potensi Minyak Jelantah Sebagai Biodiesel dan Pengaruh Katalis Serta Waktu Reaksi Terhadap Kualitas Biodiesel Melalui Proses Transesterifikasi," *Serambi Engineering*, vol. V, no. 2, 2020.

[3] R. Ramadhan, U. N. Harahap, and R. H. Nasution, "Penerapan metode Six Sigma pada pengendalian kualitas minyak goreng di PT. INNO-Wangsa Oil & Fat," *Jurnal VORTEKS*, vol. 3, no. 1, pp. 141–148, Apr. 2022, doi: 10.54123/vorteks.v3i1.137.

[4] F. Ahmad, "SIX SIGMA DMAIC SEBAGAI METODE PENGENDALIAN KUALITAS PRODUK KURSI PADA UKM," *JISI : JURNAL INTEGRASI SISTEM INDUSTRI VOLUME*, vol. 6, 2019, doi: 10.24853/jisi.6.1.11-17.

[5] J. Hasil, P. Dan, K. Ilmiah, N. Nurhayani, S. R. Putri, and A. Darmawan, "Analisis Pengendalian Kualitas Produk Outsole Sepatu Casual menggunakan Metode Six Sigma DMAIC dan Kaizen 6S."

[6] I. Septiawan, M. S. Ningsih, and I. Gunawan, "Analisis pengendalian kualitas pada Crude Palm Kernel Oil dengan metode Six Sigma di PT. X," *Jurnal VORTEKS*, vol. 3, no. 1, pp. 166–173, Apr. 2022, doi: 10.54123/vorteks.v3i1.153.

[7] A. Rahman and S. Perdana, "Analisis Perbaikan Kualitas Produk Carton Box di PT XYZ Dengan Metode DMAIC dan FMEA," *Jurnal Optimasi Teknik Industri*, 2021.

[8] B. R. Siwi, S. Nugroho, and W. P. St, "APLIKASI SIX SIGMA DMAIC DAN KAIZEN SEBAGAI METODE PENGENDALIAN DAN PERBAIKAN KUALITAS PRODUK PT. SARANDI KARYA NUGRAHA."

[9] Y. Ngatilah, U. Pembangunan Nasional, J. Timur Surabaya, C. Pujiastuti, I. Sains, and T. Akprind Yogyakarta Yogyakarta, "Use of Six Sigma and Kaizen Methods to Reduce Concrete Iron Defects (Case Study of PT. Hanil Jaya Steel)," 2018.

[10] L. J. Cheng, "Implementing Six Sigma within Kaizen events, the experience of AIDC in Taiwan," *TQM Journal*, vol. 30, no. 1, pp. 43–53, 2018, doi: 10.1108/TQM-02-2017-0017.

[11] P. L. King, "KAIZEN EVENTS AS LEAN SIX SIGMA PROJECTS RAPID IMPROVEMENT WITH STRUCTURE AND DISCIPLINE," 2010.

[12] M. Ramadhan, "Analisis Pengendalian Kualitas Produksi Untuk Mengurangi Cacat Pada Produk Sepatu Menggunakan Metode Six Sigma dan Kaizen," *MATRIK : Jurnal Manajemen & Teknik Industri-Produksi*, vol. XXII, no. 1, pp. 2621–8933, 2021, doi: 10.350587/Matrik.

[13] A. Widodo and D. Soediantono, "Benefits of the Six Sigma Method (DMAIC) and Implementation Suggestion in the Defense Industry: A Literature Review," *INTERNATIONAL JOURNAL OF SOCIAL AND MANAGEMENT STUDIES (IJOSMAS)*, vol. 3, no. 3, 2022.

[14] A. D. Wardana and Nina Mahbubah, "Integrating Seven Tools and Kaizen Approach in Evaluating Defects on Tofu Production Process," *Jurnal E-Komtek (Elektro-Komputer-Teknik)*, vol. 6, no. 1, pp. 101–113, Jun. 2022, doi: 10.37339/e-komtek.v6i1.879.

[15] R. Saputri, P. Vitasari, E. Adriantantri, and P. Studi Teknik Industri S-, "IDENTIFIKASI TIMBULNYA PRODUK CACAT DENGAN METODE CTQ DAN DPMO PADA HOME INDUSTRY KERIPIK TEMPE SARI RASA," *Jurnal Mahasiswa Teknik Industri)*, vol. 5, no. 1, 2022.

[16] M. Amerta Ivanda and H. Suliantoro, “ANALISIS PENGENDALIAN KUALITAS DENGAN METODE SIX SIGMA PADA PROSES PRODUKSI BARECORE PT. BAKTI PUTRA NUSANTARA.”
[17] A. A. Hidayat, M. Kholil, J. Haekal, N. A. Ayuni, and T. Widodo, “Lean Manufacturing Integration in Reducing the Number of Defects in the Finish Grinding Disk Brake with DMAIC and FMEA Methods in the Automotive Sub Industry Company,” *International Journal Of Scientific Advances*, vol. 2, no. 5, 2021, doi: 10.51542/ijscia.v2i5.7.
[18] H. Kartika, “| 22-32 JSTI JurnalSistem Teknik Industri *Corresponding author at:Jl Raya Meruya Selatan,” 2020.
[19] Eko Prasetyo Prayogi Sihombing, “Analisa Pengendalian Kualitas Produk Dengan Metode Six Sigma dan Analisa Kaizen di CV. Bintang Terang Medan,” *Manufaktur: Publikasi Sub Rumpun Ilmu Keteknikan Industri* , vol. 2, pp. 1–18, 2024.
[20] A. Ridwan, F. Arina, and A. Permana, “Peningkatan kualitas dan efisiensi pada proses produksi dunnage menggunakan metode lean six sigma (Studi kasus di PT. XYZ),” *Teknika: Jurnal Sains dan Teknologi*, vol. 16, no. 2, p. 186, Dec. 2020, doi: 10.36055/tjst.v16i2.9618.

Advance Sustainable Science, Engineering and Technology (ASSET)
Vol. 6, No.2, April 2024, pp. 02402021-01 ~ 02402021-08
ISSN: 2715-4211 DOI: https://doi.org/10.26877/asset.v6i2.18727

Optimizing Sample Warehouse Layout At PT BIJ: Comparative Study Of Dedicated And Class-Based Storage

Muhammad Wildan Arif*, Deny Andesta.

Teknik Industri, Universitas Muhammadiyah Gresik. Jl Sumatra 11 GKB, Gresik 61121, Indonesia

*wildanarif964@gmail.com,

Abstract. PT. BIJ is an analytical laboratory services company, therefore PT. BIJ needs a good warehouse to be able to accommodate the samples it has PT. BIJ experienced problems with sample storage. Therefore, it is necessary to improve the sample storage system to make it easier to find samples to be analyzed. Using the Dedicated Storege and Class Base Storage methods will help provide grouping for each type of sample and place the samples in the right place so that the sample travel distance is shorter. The results of the research carried out can be obtained that the initial travel distance is 262.98 meters, while the proposed dedicated storage distance is 163.19 meters and the proposed basic class storage distance is 149.95 meters, so there is the largest percentage difference of 42.89%. So it can be concluded that using the class base storage method is better than the dedicated storage method.

Keywords: Warehouse Optimization, Dedicated Storage , Class Base Storage

(Received 2024-03-25, Accepted 2024-04-11, Available Online by 2024-04-25)

1. Introduction

In the current era of development, more and more industries are developing in Indonesia, so they require support or encouragement from disciplined management to be able to compete in the industrial world. Where many factors become a reference in a business, one of which is an effective and efficient warehouse layout [1]. A warehouse is a temporary storage place for raw materials, process materials and finished materials [2]. If you just organize it without designing a good warehouse layout, you will experience difficulties in the operational process of an industry [3]. at PT. BIJ experienced problems with sample storage where samples were only placed in empty places. Therefore, it is necessary to improve the sample storage system to make it easier to find samples for analysis.

For this reason, it is necessary to arrange storage locations in the warehouse so that the warehouse becomes an effective and efficient storage place. The dedicated storage and class base storage methods will help organize products by placing a product in a good and correct storage location[4]. This placement is based on a comparison of the activities of each product with the space requirements required by that product, then the product order is obtained from largest to smallest. The aim of this method is to improve the layout of the finished product warehouse which makes it

easier to store, arrange and retrieve finished goods in the warehouse[5]. Dedicated storage or what is also known as a fixed storage location (fixed slot storage), uses a specific location or storage area for each item being stored. This is because a storage location is assigned to a specific product [6]. Class-based Storage Method This storage location rule is between the dedicated storage and randomized storage rules. Class-based storage is based on Pareto's law by taking into account the level of storage activity [7].

PT. BIJ is a company that works in the service sector, some jobs in the company will get a sample of the results of the work that has been done, therefore the company needs a room or storage warehouse to store the samples obtained from the work that has been done[8]. PT. BIJ stores samples in the warehouse for a predetermined time, therefore it requires a storage place for these samples. This company has 4 types of samples, namely Container Samples, Local Samples, Export Samples and Import Samples. In the storage warehouse, a layout that suits the needs of the samples in the warehouse has not been implemented so that storage is not optimal. The samples in the warehouse looked messy, there were several samples that were mixed together in inappropriate places and it was difficult to find the samples when taking them.

Given these conditions, companies need appropriate sample inventory management[9]. The method used to help Dedicated Storage and Class Base Storage. Dedicated Storage is a method that is often referred to as specific and permanent storage because the location for each item has been determined[10]. Meanwhile, Class Base Storage Class Based Storage is a method of storing goods by dividing items into classes based on the similarity of certain criteria such as type of material, type of movement, percentage of items and others[11]. The number of storage locations for a product must be able to meet the maximum storage space requirements for the product. The storage space required is the cumulative of the maximum storage requirements for each type of sample if more than one type of sample is to be stored[12].

2. Methods

Data collection is based on research, namely measurements, direct observation and also based on interviews with workers who work at PT BIJ. The data required for this research relates to the size of the storage warehouse, type of goods (sample), item data (sample), storage shelf capacity and material handling equipment used [13]. By using 2 methods, namely dedicated storage and class base storage to increase mileage in the PT BIJ sample warehouse. This method was chosen because it is a suitable method for use in PT BIJ sample storage warehouse and is able to group samples according to the type of sample so that it can make finding samples easier.

2.1 Throughput

Throughput is a dynamic measurement of activity or storage, which indicates the flow in storage. The term throuhput is used as a measure of the amount of storage and retrieval activity that occurs per time period[14].

$$Tj = \frac{Average\ Sample\ in}{Freight\ Capacity} + \frac{Average\ Sample\ out}{Freight\ Capacity}$$

2.2 Space Requirement

Space Requirement is the space requirement or space of each item to be stored in the warehouse, with the Space Requirement will facilitate the division of places and tell how many storage sample racks are needed to store goods[15].

$$Sj = \frac{Average\ Sample\ in\ x\ Long\ storage\ time}{Sample\ Rack\ Capacity}$$

2.3 Comparison between Throughput and Space requirements

Purposed is to divide the products or goods stored into several parts, namely high, low and medium activity. With this division, it will be easier to plan the layout in the

warehouse based on a comparison of throughput and space requirements, besides that it will make it easier to group goods according to specifications[16].

$$T/s = \frac{\text{Throughput}}{\text{Space Requirement}}$$

2.4 Class Formation

Class formation is carried out based on throughput which shows the frequency of items in the warehouse in order to obtain a position to place samples using the class base storage method. Classes are formed using the Pareto approach by calculating the percentage of interest[17].

2.5 Travel Distance Calculation Results and Comparison Results Method

is the final calculation which calculates the results of the existing mileage and calculates the results of comparing the distance traveled by the two methods used. [18]

$$\text{Travel Distance} = \frac{T}{s}\ X\ Sj\ X\ Distance$$

$$\text{Comparison Results Method} = \frac{\text{Original Layout} - \text{Proposed Layout}}{\text{Original Layout}}$$

3. Results and Discussion

Data collection

Data collection The data collection process was carried out by means of interviews and direct observation, the following data was obtained:

Table 1 Sample Date And Storage Shelf Capacity

Data Perusahaan					
No	Name	Average Incoming Sample	Time Save	Total Sample At warehouse	Capacity Sample Rack
1	Container Sample	1120 Sample	3 Bulan	3360 Sample	16800
2	Local Sample	9182 Sample	3 Bulan	27546 Sample	16800
3	Import Sample	1216 Sample	6 Bulan	7296 Sample	16800
4	Export Sample	4288 Sample	6 Bulan	25728 Sample	16800

The data above is data obtained from the company which can be known from this data. The number of samples in PT BIJ's storage warehouse is very large, therefore improvements are needed in the sample storage warehouse. By using dedicated storage and class base storage methods, we are able to perfect the existing storage system at PT BIJ. which by using this method the samples in the warehouse will be grouped according to the existing types and will make it easier for employees to find samples to be analyzed.

Data processing

Throughput Calculation

Based on the results of calculating samples entering and leaving PT BIJ's finished goods warehouse, it shows that the throughput is 84.23 times. This means that the total transfer trip activities for storage and retrieval that occurred within a month's time period was 84.23 times.

Table 2 Calculation Througput

Calculation Througput

No	Name	Average Incoming Sample	Average Samples out	Througput
1	Container Sample	1120	1400	6,72
2	Local Sample	9182	8811	47,98
3	Import Sample	1216	1792	8,02
4	Export Sample	4288	3776	21,50
	Total Througput			84,23

Calculation of Space Requirements

The sample warehouse has an area of 35 M^2 with a length of 7 M and a width 5 M which will be divided into 6 storage shelf slots. The calculation of space requirements on the storage shelf for each sample is carried out by rounding up to ensure that no samples lack storage space.

Table 3 Calculation of Space Requirements

Calculation of Space Requirements				
No	Name	Average Incoming Sample	Space Requirement Teoritis	Space Requirement Sj
1	Container Sample	1120	0,2	1
2	Local Sample	9182	1,64	2
3	Import Sample	1216	0,43	1
4	Export Sample	4288	1,53	2

Comparison between Throughput and Space requirements

After getting the results of calculating the frequency of receipts and expenditures (throughput) and slot requirements (Space Requirement), a calculation will be carried out for the ratio Tj and Sj. The sample with the largest T/S ratio value will be placed in the slot with the smallest distance, the product with the second largest T/S ratio value will be placed in the slot with the second smallest distance, and so on. Technically, this placement method aims to reduce the operator's travel distance from the I/O point to the storage rack.

Table 4 Comparative Calculations Tj And Sj

Comparison between Throughput and Space requirements				
No	Name	Sj	Througput	Results
1	Container Sample	1	6,72	6,72
2	Local Sample	2	47,98	23,99
3	Import Sample	1	8,02	8,02
4	Export Sample	2	21,50	10,75

Formation of Classes Based on Class Storage Methods

Table 5 Formation of Classes Based on Class Storage Methods

Formation of Classes Based on Class Storage Methods						
No	Sample Type	Througput	Percentage Througput	Amount Sample	Percentage Sample	Code
1	Local Sample	47,98	57%	9182	58%	A
2	Export Sample	21,50	26%	4288	27%	B
3	Import Sample	8,02	10%	1216	8%	C
4	Container Sample	6,72	8%	1120	7%	D

Travel distance calculation

The table below shows the distance between the exit door and the sample storage shelf which is calculated directly when in the sample storage warehouse. And the proposed storage rack spacing table, taking into account the comparison results of throughput and storage, results are obtained as in the table.

Table 5 Original Mileage

Original Mileage					
No	Name	T/s	Sj	Distance	Mileage
1	Container Sample	6,72	1	1 Meter	6,72
2	Local Sample	23,99	2	4 Meter	177,53
3	Import Sample	8,02	1	4 Meter	38,10
4	Export Sample	10,75	2	1,5 Meter	45,16
Total Mileages					267,51

Table 6 Proposed Mileage of Dedicated Storage

Proposed Dedicated Storage Mileage					
No	Name	T/s	Sj	Distance	Mileage
1	Container Sample	6,72	1	4 Meter	26,88
2	Local Sample	23,99	2	1,5 Meter	71,97
3	Import Sample	8,02	1	4 Meter	32,09
4	Export Sample	10,75	2	1,5 Meter	32,26
Total Mileages					156,10

Table 7 Proposed Mileage of Class base Storage

Proposed Base Class Storage Mileage					
No	Name	T/s	Sj	Distance	Mileage
1	Container Sample	6,72	1	4 Meter	26,88
2	Local Sample	23,99	2	1 Meter	47,98
3	Import Sample	8,02	1	4 Meter	32,09
4	Export Sample	10,75	2	2 Meter	43,01
Total Mileages					149,95

From the results of the comparison table above, it can be concluded that the two methods are able to provide closer distance changes and can shorten the activity time that occurs in the warehouse and save energy expended. Using this method can make it easier for employees to find and place similar samples on designated storage shelves.

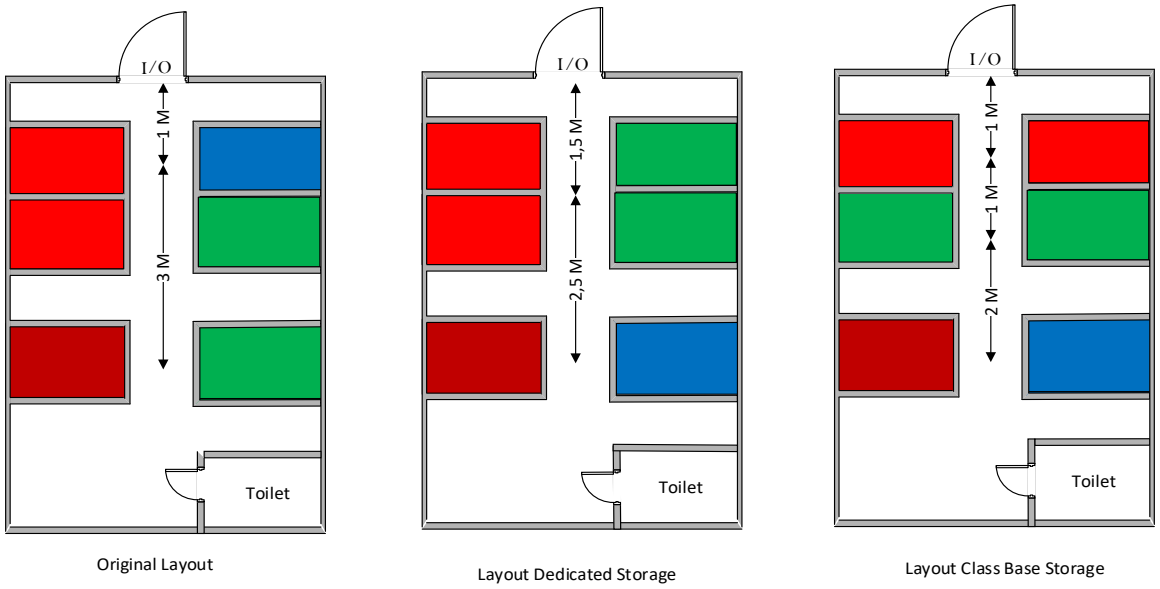


Figure 1. Original And Proposed Warehouse Floor Plan Images

Information :

= Local Sample Place

= Import Sample Place

= Container Sample Place

= Export Sample Place

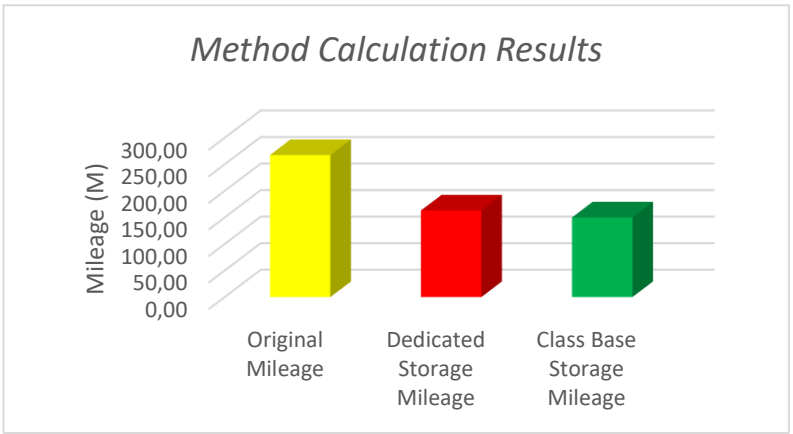


Figure 2 Method Calculation Results

Table 8 Method Calculation Results

Method Calculation Results			
Layout	Total Distance	Difference	Percentage Difference
Original Layout	262,98		
Usulan Dedicated Storage	163,19	99,79	37,95%
Usulan Class Base Storage	149,95	113,03	42,98%

4. Conclusion

In the research that has been completed, several results have been achieved, namely the initial total distance traveled was 262.98 meters. The distance traveled by the proposed Dedicated Storage method is 163.19 meters. The distance traveled by the proposed Class Base Storage method is 149.95 meters so there is a decrease of 37.95% for the Dedicated Storage method and 42.98% for the Class Base Storage method so that the comparison between the two methods is has the best improvement is the Class Base Storage method.

The suggestion given is that by using the class base storage method, PT BIJ's sample storage warehouse is able to shorten the sampling distance in the warehouse and also group samples according to their respective types and make it easier to find certain types of samples to be taken for analysis.

References

[1] Indrawan, Andi Steven. "Perbaikan Tata Letak Gudang Distribusi dengan Data mining, Dedicated Storage dan Multi-product Slot Allocation." Jurnal Teknik Industri 12.1 (2022): 9-20.

[2] Suwarno, Basuki Arianto, and Karel L. Mandagie. "Perancangan Tata Letak Gudang Produk Jadi Cat Dengan Metode Dedicated Storage Di Pt. Akzonobel Car Refinishes Indonesia." Jurnal Teknik Industri 8.2 (2021).

[3] Prasetyo, Yuyut Tri, and Ahmad Fatih Fudhla. "Perbaikan Tata Letak Fasilitas Gudang Dengan Pendekatan Dedicated Storage Pada Gudang Distribusi Barang Jadi Industri Makanan Ringan." Jurnal Teknik Industri: Jurnal Hasil Penelitian Dan Karya Ilmiah Dalam Bidang Teknik Industri 7.1 (2021): 1-6.

[4] Anam, Muhammad Syaiful, Moch Nuruddin, and Hidayat Hidayat. "Usulan Perbaikan Layout Area Gudang Kering Logistik Menggunakan Metode Dedicated Storage di PT. Starfood International." Jurnal Serambi Engineering 8.4 (2023).

[5] Sarathamani, Anamika Palavesam, and Anil Kumar. "Fuzzy machine learning model for class-based flood damage assessment from planetscope temporal data." Journal of Applied Remote Sensing 18.1 (2024): 014523-014523.

[6] Sitorus, Helena, Rudianto Rudianto, and Meriastuti Ginting. "Perbaikan Tata Letak Gudang dengan Metode Dedicated Storage dan Class Based Storage serta Optimasi Alokasi Pekerjaan Material Handling di PT. Dua Kuda Indonesia." Jurnal Kajian Teknik Mesin 5.2 (2020): 87-98.

[7] Irman, Ade, Y. Muharni, and Andri Yusuf. "Design of warehouse model with dedicated policy to minimize total travel costs: a case study in a construction workshop." IOP Conference Series: Materials Science and Engineering. Vol. 909. No. 1. IOP Publishing, 2020.

[8] Anto, C., et al. "Redesign production layout using dedicated storage method: case study of PT. Solo Grafika Utama." IOP Conference Series: Materials Science and Engineering. Vol. 943. No. 1. IOP Publishing, 2020.

[9] Shafira, Sasya. "PERANCANGAN ULANG TATA LETAK GUDANG BAHAN BAKU DENGAN METODE DEDICATED STORAGE UNTUK MEMINIMALKAN JARAK PEMINDAHAN BAHAN DI PT. X" . Diss. UPN VETERAN JATIM, 2022.

[10] El Islami, Nur, and Helshy Azzahra. "ANALISIS PERANCANGAN TATA LETAK GUDANG CURAH MENGGUNAKAN METODE DEDICATED STORAGE (STUDI KASUS: PT KRAKATAU JASA LOGISTIK)." (2022)
[11] Utama, Hendrawan Taksaka. Analisis Penempatan Barang Menggunakan Metode Dedicated Storage Pada Gudang Bahan Baku PT Agrofarm Nusa Raya. Diss. Politeknik APP Jakarta, 2022.
[12] Ferdiansyah, Fikar Rizki. "ANALISIS RELAYOUT PADA GUDANG BARANG MENGGUNAKAN METODE DEDICATED STORAGE (Studi Kasus pada Home Industri Alika Collection Soreang Kabupaten Bandung)". Diss. STIE Ekuitas, 2023.
[13] Utari, Winda Ayuningtyas. "REDUKSI AREA TRANSIT PART OVERFLOW DENGAN PERBANDINGAN METODE DEDICATED STORAGE DAN METODE CLASS BASED STORAGE". Diss. Universitas Pembangunan Nasional Veteran Jakarta, 2020.
[14] Alfian, Achmad, and Shendy Pratama. "Perancangan Tata Letak Warehouse Produk menggunakan metode dedicated storage di PT Nutrifood Indonesia." Jurnal Ilmiah Teknik Industri 10.1 (2022): 77-85.
[15] Rosihin, Rosihin, Dadi Cahyadi, and Supriyadi Supriyadi. "Analisa Perbaikan Tata Letak Gudang Coil dengan Metode Class Based Storage." Jurnal INTECH Teknik Industri Universitas Serang Raya 7.2 (2021): 166-172.
[16] Isnaeni, Nadila Safira, and Novie Susanto. "Penerapan Metode Class Based Storage Untuk Perbaikan Tata Letak Gudang Barang Jadi (Studi Kasus Gudang Barang Jadi K PT Hartono Istana Teknologi)." Industrial Engineering Online Journal 10.3 (2021).
[17] Kulsum, Kulsum, Yusraini Muharni, and Al-Atur Akbar Felayani. "Usulan pengoptimalan tata letak gudang W12 menggunakan kebijakan dedicated storage dengan penerapan simulasi (Studi kasus: PT. XYZ)." Teknika: Jurnal Sains dan Teknologi 16.2 (2020): 285-292.
[18] Novrianto, Tsany Farras, Deny Andesta, and Moh Jufriyanto. "Usulan Tata Letak Pada Gudang Sparepart dengan Kebijakan Class Based Storage Berdasarkan Analisis ABC di PT. XYZ." G-Tech: Jurnal Teknologi Terapan 8.2 (2024): 975-984

Advance Sustainable Science, Engineering and Technology (ASSET)
Vol. 6, No.2, April 2024, pp. 02402022-01 ~ 02402022-09
ISSN: 2715-4211 DOI: https://doi.org/10.26877/asset.v6i2.18798

The Analysis of Habitat Suitability for Macaca Tonkeana in the Pangi Binangga Natural Reserve of West Toboli Village Using SIG and PCA Approach

Siti Susianti*, Hasriani Muis, Misrah, Akhbar, Ida Arianingsih, Rhamdhani Fitrah Baharuddin, Amati Eltriman Hulu

Faculty of Forestry, Tadulako University, Tadulako, Jl. Soekarno Hatta Km. 9 Palu 94119, Central Sulawesi, Indonesia

*susianti2571@gmail.com

Abstract. The boti monkey, also known by its scientific name Macaca tonkeana, is a type of endemic primate found on Sulawesi Island. This study aims to determine the suitability of Macaca tonkeana habitat by using spatial analysis methods to determine the most suitable areas for survival. The study used a sample of locations conducted by observation method. The data collected was then spatially analysed by overlaying, classing, weighting, and scoring. PCA analysis was used to classify variables and assign weights to each parameter. Based on the habitat suitability model developed, the study sites were classified into three suitability levels: low, medium and high. Low suitability with an area of 9.475621 ha (0.374%). Medium suitability with an area of 461.777729 ha (18.184 %). High suitability with an area of 2065.896076 ha (81.442%). The model validation showed that most of the validation points were located within the high suitability class, with 92.59% of the total validation points. This indicates that the study site is likely to be a highly suitable habitat for Boti Monkeys.

Keywords: Habitat Suitability, Macaca Tonkeana, SIG, Spatial Analysis

(Received 2024-03-30, Accepted 2024-04-11, Available Online by 2024-04-25)

1. Introduction

Central Sulawesi is among the provinces in Indonesia in the center of Wallacea, a region with a unique wealth of biological natural resources, which includes abundant endemic flora and fauna [1]. Flora consists of all types of plants that can be found in Central Sulawesi Province, while fauna refers to the various types of animals that exist in the region [2].

Central Sulawesi, precisely in Parigi Moutong Regency, has a nature reserve, namely the Pangi Binangga Nature Reserve, where there is flora and fauna [3]. According to Law Number 5 of 1990 concerning the Conservation of Biological Natural Resources and Ecosystems, Nature Reserves (CA) are protected natural areas because they have unique plants, animals, and ecosystems or specific ecosystems that require natural protection. The determination of CA Pangi Binangga is based on the Decree of the Minister of Forestry and Plantation Number 399/Kpts-2/1998 with an area of more than 6,000 hectares, administratively located in Parigi Moutong district, Central Sulawesi Province.

Macaca tonkeana is one of seven endemic monkey species that can be found in the central region of Sulawesi [4]. One of them is along the Trans Palu-Parigi road, which is located in the Pangi Binangga Protected Forest and Nature Reserve. This area is the natural habitat (home range) of Macaca tonkeana.

The boti monkey, also known by its scientific name Macaca tonkeana, is a type of endemic primate found on Sulawesi Island [5]. The Boti Monkey (Macaca Tonkeana) in West Toboli Village is a protected animal, but currently, its existence is threatened, so there is a need for data regarding the suitability of the black monkey's (Macaca Tonkeana) habitat. One of the main reasons why the existence of the boy monkey (Macaca tonkeana) is threatened is the destruction of its habitat due to human activities in forest exploitation, natural disasters, and lack of food availability [6].

Spatial research combines GIS, satellite imagery, and aerial photos to evaluate habitat factors for Macaca tonkeana conservation, aiming to understand parameter impacts and develop a habitat suitability model for West Toboli Village's Pangi Binangga Nature Reserve. [7]. This research aims to identify optimal areas for the black monkey's survival, recognizing the significance of distribution data and spatial analyses for conservation efforts.[6].

2. Methods

2.1 Research Location and Time

This research was carried out in the Pangi Binangga Nature Reserve area, West Toboli Village, Parigi Moutong Regency, Central Sulawesi. This research was carried out from September to November 2023.

2.2. Data Collection and Analyses

The research was conducted by observing and sampling locations, collecting precise coordinates of monkey nests using GPS. The collected data was then analysed using spatial analysis techniques such as overlay, class division, weighting, and scoring, which are essential for identifying spatial models and relationships between variables [8]. Data were collected using a purposive method, by searching for nests and signs of boti monkeys, selected based on information from local communities and conservation officers. Nest sites were selected by considering the nightly nesting habits of boti monkeys. Sites were also selected with respect to distance from roads. Boti monkey nests and tracks were identified and their coordinates recorded using GPS.

The stages for spatial analysis of habitat suitability are as follows:

1. Collect data based on each function to create a distribution map of boti monkeys. The types of data required include coordinate points (GPS points) and Landsat 8 images.
2. Spatial Distribution Mapping. Analysis of the spatial distribution of boti monkeys was carried out by utilizing all coordinate points obtained based on their location [9]
3. Habitat suitability parameters include distance from settlements, roads, rivers, slope, height, and NDVI for land cover.[10]
4. Distance Maps are created based on predetermined variables, then buffering the parameters used via ArcGIS 10.8 [11].
5. Making vegetation density maps. Normalized Difference Vegetation Index (NDVI) is used to determine vegetation density[12].
6. The use of PCA analysis aims to identify the factors that have the most significant impact [13]. The PCA analysis was carried out using XLSTAT software. The PCA results are used to determine the weight of each habitat factor and for spatial analysis to produce the following equation ;

$$Y = aFK1 + bFk2 + cFk3 + dFk4 + eFk5 + fFk6 + gFk7 \quad (1)$$

Information:

Y : Total Habitat Suitability Value a-f: Weight Value of Each Variable
Fk1: Distance Factor From River Fk2: Distance Factor From Road
Fk3: Distance Factor from Settlements Fk4: NDVI Factor
Fk5 : Slope Factor Fk6: location Height Factor

7. In spatial analysis, important factors include the location of macaca tonkeana distribution points, which are considered in terms of distance from rivers, road networks, and settlements, as well as NDVI values [14].

The Mathematical Model used is:

- The interval value of the habitat suitability classification score is determined based on the distribution of pixel values:

$$Interval = \frac{Smax - Smin}{K} \quad (2)$$

Information:

Smax : Highest Pixel Value Smin : Lowest Pixel Value

K : There are many classifications of habitat suitability

- Validation value of habitat suitability classification:

$$Validation = (n - N)\ 100\ \% \quad (3)$$

3. Results and Discussion

Spatial habitat suitability modeling is essential in understanding and protecting sustainability. In creating a spatial model of forest suitability, several data are needed, including habitat suitability parameters, PCA data analysis, and spatial data analysis.

3.1. Habitat Suitability Parameters

Six parameters - settlement, roads, rivers, elevation, slope, and NDVI - are analyzed to form a habitat suitability model for boti monkeys..

3.1.1 Distribution of Nests Based on Distance from the River

The results of the nest distribution buffer parameters (n) and area size based on distance from the river are presented in Table 1.

Table 1. Nest distribution (n) and area area based on distance class from the river

Class	N	Nest (%)	Area (ha)	Area Percentage (%)	Nest Density/ha
0 - 2000	54	100%	2527,673402	100%	699,63
Total	54	100%	2527,673402	100%	

Table 1 displays nest distribution and area relative to river distance. Within 0-2000 meters, 54 nests cover 2527.673402 hectares, with a nest density of 699.63 per hectare.

3.1.2 Distribution of Nests Based on Distance from Settlements

Settlement is a factor that influences the habitat of boti monkeys. The results of the nest distribution buffer parameters (n) and area size based on distance from the river are presented in Table 2.

Table 2. Nest distribution (n) and area based on distance class from settlement

Class	N	Nest (%)	Area (ha)	Area Percentage (%)	Nest Density/ha
0 – 1000	0	0%	331,347126	13%	0
1000 – 3000	14	26%	700,669164	28%	1998,31
3000 – 5000	40	74%	815,129139	32%	4906,75
5000 – 7000	0	0%	671,091259	27%	0
> 8000	0	0%	9,436714	0%	0
Total	54	100%	2527,673402	100%	

Table 2 presents nest distribution and area by distance from the reference point to the settlement. Majority of nests were within 1000-5000 meters, with densest concentration (4906.75/ha) at 3000-5000 meters, suggesting optimal habitat for boti monkeys.

3.1.3 Nest distribution based on distance from the road

The results of the nest distribution buffer parameters (n) and area size based on distance from the road are presented in Table 3.

Table 3. Nest distribution (n) and area based on distance class from road

Class	N	Nest (%)	Area (ha)	Area Percentage (%)	Nest Density/ha
0 – 1000	54	100%	1703,72846	67%	2297,05

1000 – 3000	0	0%	815,535556	32%	0
3000 – 5000	0	0%	8,409386	0%	0
Total	54	100%	2527,673402	100%	

Table 3 shows nest distribution and area by road distance. In the 0-1000 meter class, 54 nests cover 1703.72846 hectares with a nest density of 2297.05 nests/ha. Subsequent classes have zero nests.

3.1.4 Distribution of Boti Monkey Nests Based on NDVI

Vegetation density is a factor that influences the habitat of boti monkeys. The results of nest distribution (n) and area size based on NDVI are presented in Table 4.

Table 4. Nest distribution (n) and area based on NDVI

Class	N	Nest (%)	Area (ha)	Area Percentage (%)	Nest Density/ha
0-0,3	3	6%	348,693824	14%	0,86
0,3-0,4	10	19%	884,071802	35%	1,13
0,4-0,5	37	69%	1098,10306	43%	3,37
>0,5	4	7%	196,804716	8%	2,03
Total	54	100%	2527,673402	100%	

Table 4 illustrates nest distribution by NDVI, reflecting boti monkey habitat. Classes range from 0 to >0.5, correlating with vegetation. Class 0.4-0.5 harbored most nests (69%), covering 43% of the area. Higher NDVI classes, though fewer nests, maintained high nest density, emphasizing NDVI's significance in nest distribution.

3.1.5 Nest Distribution Based on Height

Height is a factor that influences the habitat of boti monkeys. The results of nest distribution (n) and area size based on height are presented in Table 5.

Table 5. Nest distribution (n) and area based on height

Class	N	Nest (%)	Area (ha)	Area Percentage (%)	Nest Density/ha
0-300	4	7%	792,811956	31%	5,05
300-400	1	2%	282,42392	11%	3,54
400-500	5	9%	308,288903	12%	16,16
500-750	42	78%	692,47269	27%	60,65
>750	2	4%	451,675933	18%	4,43
Total	54	100%	2527,673402	100%	

Table 5 displays nest distribution across five altitude classes, from 0-300 meters to over 750 meters. Most nests (78%) were found at 500-750 meters, suggesting altitude's importance in boti monkey nest distribution.

3.1.6 Nest Distribution Based on Slope

The slope is a factor that influences the habitat of boti monkeys. The results of nest distribution (n) and area size based on slope are presented in Table 6.

Table 6. Nest distribution (n) and area area based on slope

Class	n	Nest (%)	Area (ha)	Area Percentage (%)	Nest Density/ha
0-8	23	43	0,004756	0	4831679,01
8-15	13	24	799,075532	32	1625,78
15-25	16	30	806,72522	32	1984,59

25-40	0	0	738,191813	29	0
>40	2	4	161,047063	6	1242,33
Jumlah	54	100	2505,044384	100	

Table 6 depicts orangutan nest density by slope class. The 0-8 degree class exhibits the highest density (23 nests, 43%) in a small area (0.004756 ha). Slopes of 0-25 degrees contain most nests, emphasizing slope's role in nest distribution.

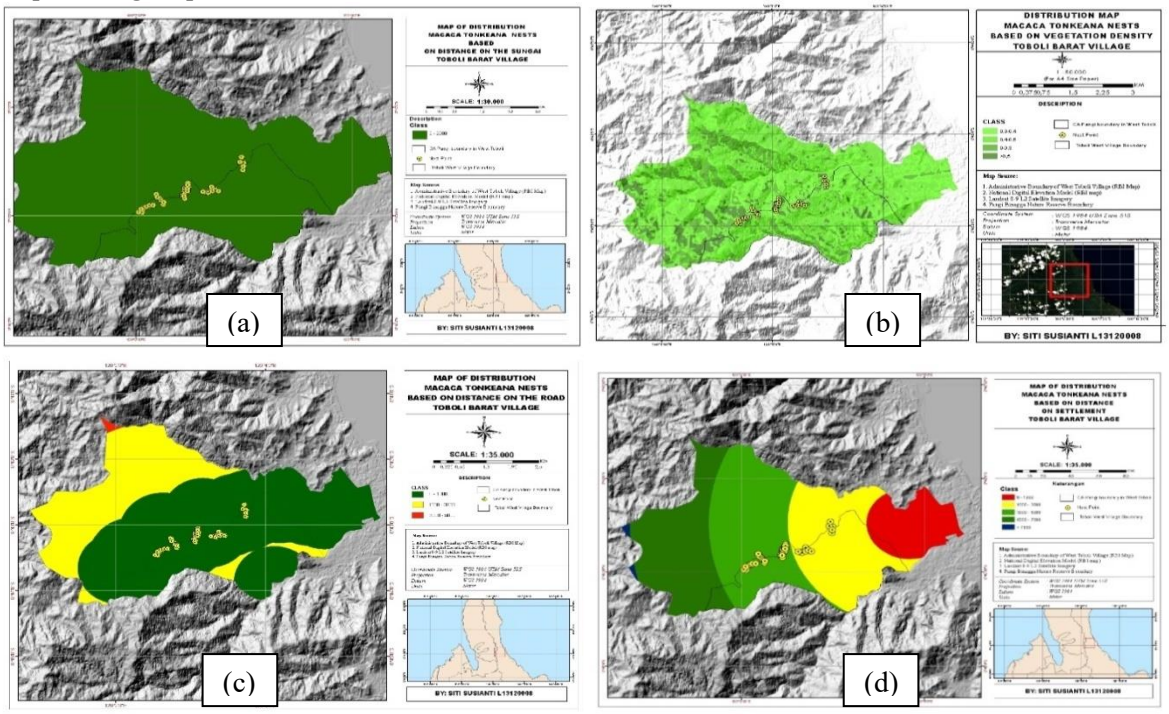


Figure 1. (a). map of macaca tonkeana nest distribution distance from rivers (b). map of macaca tonkeana nest distribution NDVI (c). map of macaca tonkeana nest distribution distance from roads (d). map of macaca tonkeana nest distribution distance from settlements.

3.2 Habitat Suitability Model Data Analysis Using PCA

PCA identifies relationships between variables to create new independent variables. Previous studies [15] have indicated that PCA can be used to assign scores or weights to variables in predicting the habitat of a species. Therefore, in this study, PCA was utilized to classify variables and determine scores or weights for each group of variables or factors (PC) generated. Meanwhile, to ascertain the weights of each tested variable, regression analysis was conducted using the factor scores of each PC against all variables that formed it. PCA analysis and regression analysis were performed using the XLSTAT software.

From the results of the PCA analysis involving 6 principal components, 3 principal components were found to be significant in the total variation, which are described in Table 7. The usable and representative principal component is the third principal component with a cumulative diversity value of 76.1%. The cumulative diversity value is considered to represent the total diversity of the data, because the cumulative diversity lies between 70%-80%.

Table 7. Total Diversity of Principal Components

Main component	**Characteristic Roots**		
	Total	**Diversity (%)**	**Cumulative diversity (%)**
River	3,908	78,158	78,158

Settlement	4,415	88,309	88,309
Road	4,415	88,309	88,309
NDVI	4,419	88,376	88,376
Height	3,696	73,92	73,92
Slope	3,328	66,567	66,567

Data from Table 8 analysis is utilized to establish weights for each parameter, showcasing the relationship between the six influencing parameters for boti monkey habitat suitability.

Table 8. PCA Habitat Characteristic Vector

Parameter	Main component		
	1	**2**	**3**
River	5,752	5,752	5,756
Settlement	6,499	6,499	6,504
Road	6,499	6,499	6,504
NDVI	6,504	6,504	6,509
Height	5,440	5,440	5,444
Slope	4,899	4,899	4,902

The weights of each variable to obtain the habitat suitability model for orangutans are obtained from the PCA feature vector values of each variable with the highest positive values towards the principal component produced. The results above indicate that variables such as river, road, settlement, NDVI, elevation, and slope have a high positive relationship with the third principal component. Thus, the magnitude of the weights of each variable is presented in the table below.

Hasil dari Tabel 8 menunjukkan bahwa setiap variabel dikategorikan dalam komponen utama. Sehingga Keenam parameter tersebut termasuk ke dalam kelompok komponen tiga, sehingga pembobotannya disesuaikan. Nilai koefisien untuk sungai, pemukiman, jalan, NDVI, ketinggian, dan kemiringan lereng masing-masing adalah 5,756, 6,504, 6,504, 6,509, 5,444, dan 4,902.

Information regarding the scores used in the habitat suitability analysis, where each parameter is given a score ranging from 1 to 5 to determine habitat suitability, has a different range of scores. Where Score 1 represents the worst condition for Macaca tonkeana, Score 2 represents poor condition, Score 3 represents fair condition, Score 4 represents good condition, and Score 5 represents the best condition for Macaca tonkeana. Settlement parameters have scores that vary from 1 to 5, with higher scores indicating better habitat conditions. In contrast, scores for parameters such as roads, rivers, slopes, elevation and NDVI also range from 1 to 5, but the variability depends on the specific category or range of values. Determining these values is critical to understanding the level of habitat suitability and provides the basis for appropriate conservation measures.

From the calculations carried out using XLSTAT for each variable, an equation can be formulated for the boti monkey habitat suitability model as follows:

Y = 6,504jln + 6,504pkm + 5,756sng + 6,509ndvi + 5,444elev + 4,902slope

This equation highlights that vegetation cover (NDVI), roads (jln), and settlements (pkm) have the highest coefficients (weights) compared to other variables. The variables river (sng) and height (elev) are next in order, followed by the slope (slope), which has the most negligible weight. Thus, the weight results of each parameter are used to determine the habitat suitability value from the overlay process.

3.3 Habitat Suitability for Boti Monkeys

The constructed fit equation yields maximum and minimum values, ranging from 57.18 to 160.66. These values are divided into three classes based on their differences: the smallest range indicates low suitability, the middle range indicates medium suitability, and the largest range indicates high suitability.

$$Interval = \frac{160.66 - 57.18}{3} = 14.16$$

The suitability of boti monkey habitat in the study site is classified into three levels, namely low, medium, and high suitability, which are described in Table 10. In addition, a map of boti monkey habitat suitability is also presented in Figure 2. The high suitability level reached 81.72%. Locations with high suitability are mostly in the Pangi Binangga nature reserve area, while locations with low suitability tend to be in settlements. The area of West Toboli Village, which is mostly highland and hilly, is mostly inhabited by people who work as farmers. Areas with high suitability generally have good land cover and there are many fruit trees that are a source of food for boti monkeys.

According to [16], monkeys in Sulawesi tend to choose large trees with many branches, allowing all members to rest. They also prefer the edges of tree canopies because there are usually parts of the plants they consume, such as fruit, flowers, and young leaf shoots. The characteristics of trees used as resting places by monkeys are large trees with many branches and dense leaves.

Table 9. Habitat Suitability Index Value Habitat Area for Boti Monkeys

Interval	Area (ha)	Percentage	Conformity Classification
5,444 - 57,18267	9,475621	0,374	Low
57,18267 - 108,9213	461,777729	18,184	Medium
108,9213 - 160,66	2065,896076	81,442	High
Total	2537,149426	100	

The medium suitability level, covering 18.184% (461.777729 ha) in light green on Figure 2, lies predominantly amidst settlements and agricultural plots. The region's contoured topography, with hills, valleys, and diverse elevations, contributes to its medium suitability status. Low suitability, comprising 0.374% (9.475621 ha), marked in red on Figure 10, is situated within residential zones and areas approximately 300 meters above sea level.

The existence of conservation areas must be maintained due to their strategic role as life support systems, biodiversity protection, and sustainable utilization of natural resources. In maintaining the existence of potential conservation areas, one of the applied management concepts is to exclude all community activities from the conservation areas [16]. From the habitat suitability data, it can be used to conserve the Macaca Tonkeana species to enhance their existence and protection. Several factors influencing habitat suitability include roads, rivers, settlements, NDVI, elevation, and slope.

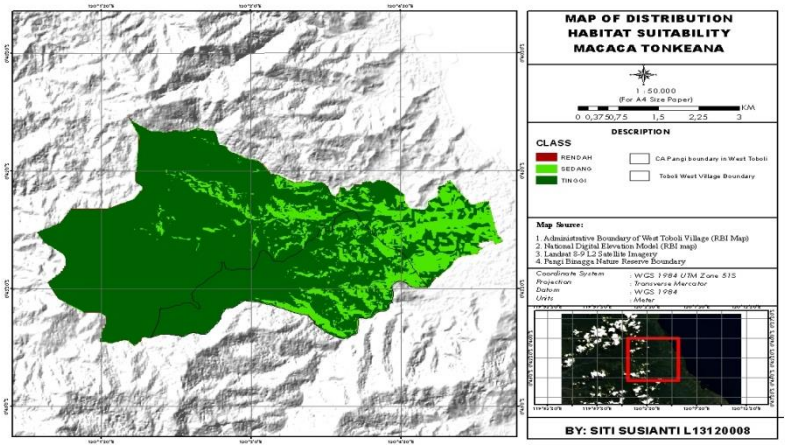


Figure 2. Boti Monkey Habitat Suitability Map

3.4 Validasi

Validation is conducted to test the accuracy of the prepared model using specially prepared validation data [14]. At this stage, fifty-four observation points (30% of the total observation data) were used. From the results of this combination, the following data is obtained in Table 10:

Table 10. Habitat suitability model validation results

Conformity Class	Number of Points	Percentage
Low	0	0
Medium	4	7,41
High	50	92.59
Total	54	100

Table 10 displays validation outcomes of the boti monkey habitat model, categorized into low, medium, and high suitability. None fell in the low class, while 4 (4%) were medium, and 50 (92.59%) were high. This suggests the research area is highly suitable for boti monkeys, affirming the model's accuracy.

4. Conclusion

The study concluded that various factors such as distance from rivers, settlements, roads, NDVI (Normalised Difference Vegetation Index), altitude, and slope affect the habitat suitability of Boti Monkeys. PCA analysis assigned weights to each parameter. The study sites were classified into low, medium, and high suitability levels, with low suitability mainly in residential areas, covering 9.475621 ha. Medium suitability was generally found among settlements, covering 461.777729 ha, while high suitability was dominated by the Pangi Binangga nature reserve, covering 2065.896076 ha. Model validation showed that the study site is a suitable habitat for Macaca tonkeana.

Acknowledgment

Thanks to the Faculty of Forestry, Tadulako University for supporting the research through the MBKM Mandiri programme. Thanks also to the Central Sulawesi Natural Resources Conservation Centre for helping and facilitating this research. Their co-operation and support means a lot to us.

References

[1] A. R. H. Bisjoe, “Kawasan Wallacea Dan Implikasinya Bagi Penelitian Integratif Lingkungan Hidup Dan Kehutanan Achmad Rizal H. Bisjoe,” Pp. 141–148.

[2] N. Hadi, N. S. Ainy, And L. Sjahfirdi,"Konservasi Dan Perlindungan Keanekaragaman Hayati : Menahan Laju Kepunahan Dan Ancaman Utama Hidupan Liar Di Indonesia,” Vol.13,No.1, 2023.

[3] P.A.R. Ramadanilpitopang, Reza Rizaldi, M.Fajri Ramadhan,“Plant Species Of Zingiberaceae In The Pangi Binangga Nature Reserve, Central Sulawesi),”Vol.11,No.28,Pp.19–27,2021.

[4] M. Trianto, N. Herjayanti, And M. D. Kisman, “Public Perception On The Existence Of Macaca Hecki And Macaca Tonkeana In Protected Forest And Pangi Binangga Nature Reserve Of Central Sulawesi” No. January 2024, 2021, Doi: 10.29303/Jbt.V21i2.2769.

[5] M. S. I. Gunawan, A. Annawaty, And ..., “Distribusi Kelompok Macaca Hecki (Matschie, 1901) Dan Macaca Tonkeana (Meyer, 1899) Di Hutan Lindung Dan Cagar Alam Pangi Binangga Sulawesi Tengah,” ... *J. Sci.* ..., Vol. 7, No. 2, Pp. 219–225, 2018.

[6] M. Syaputra, “(Ex Situ Conservation Strategies Of The Yellow Crested Cockatoo (Cacatua Sulphurea Occidentalis) At Conservation Unit,” Vol. 19, 2022.

[7] H. Prayogo And S. Pascasarjana, “Pendugaan Kesesuaian Habitat Orangutan Kalimantan (Pongo Pygmaeus Pygmaeus) Dan Karakteristik Genetiknya Di,” 2014.

[8] D. E. W. Arif Mohammad Siddiq1, H. Sulistiyowati, R. Wimbaningrum, R. Setiawan, And D. Supriadi, “Habitat Characteristics Of Long-Tailed Macaque In Kucur Resort At Alas Purwo National Park,” Vol. 10, No. 2, Pp. 94–100, 2022,

[9] A. E. Hulu, H. Muis, And S. D. Massiri, “Spatial Analysis Of Water Infiltration Potential In The Miu Watershed Of Sigi Regency,” Vol. 5, No. 2, Pp. 1–10, 2023.

[10] D. Kustanto And D. I. K. Bukittinggi, “Analisis Geospasial Sebaran Stunting Di Kota Bukittinggi,” No. March, 2022, Doi: 10.35730/Jk.V13i1.576.

[11] S. Redjeki *Et Al.*, “Spatial And Temporal Distribution Of Blue Swimming Crab Larvae (Portunus Pelagicus) In The Conservation Area Of Betahwalang , Central Java,” Vol. 25, No. December, Pp. 173–178, 2020,

[12] A. E. Hulu *Et Al.*, "Peningkatan Kompetensi Mahasiswa Kehutanan Melalui Pelatihan Penggunaan Arcgis," Vol. 1, No. 6, Pp. 854–860, 2023.
[13] N. Istiqamah, H. Muis, S.D. Massiri, And R. Fitrah, "Spatial Analysis of Bobo Village Community Accessibility Level in the Lore Lindu National Park Area," Vol.6,No.2,Pp.1–9, 2024.
[14] A. Cahyana And J. H. Dan L. B. Prasetyo, "Pemodelan Spasial Kesesuaian Habitat Elang Jawa Di Taman Nasional Gunung Halimun-Salak," Vol. 20, No. 3, Pp. 211–219, 2016.
[15] D. Pombu And A. M. , Elhayat Labiro, "Studi Habitat Monyet Boti (Macaca Tonkeana) Di Hutan Lindung Desa Sangginora Kecamatan Poso Pesisir Selatan Kabupaten Poso," Pp. 25–32, 2014.
[16] E. Setiawan, "Pengelolaan Dan Konservasi Satwa Berbasis Kearifan Lokal Di Taman Nasional Alas Purwo," *J. Sosiol. Dialekt. Sos.*, Vol. 8, No. 2, P. 113, 2022, Doi: 10.29103/Jsds.V8i2.8660.

Advance Sustainable Science, Engineering and Technology (ASSET)
Vol. 6, No.2, April 2024, pp. 02402023-01 ~ 02402023-14
ISSN: 2715-4211 DOI: https://doi.org/10.26877/asset.v6i2.18610

Sustainability Strategies of Traditional Vannamei Shrimp Cultivation in East Java: A Case Study in Kudu Hamlet, Lamongan District

Dwi Sugiarto*, Farikhah

Aquaculture and Fisheries Study Program, Faculty of Agriculture, Universitas Muhammadiyah Gresik

*ugi06.05.2002@gmail.com

Abstrak. The research investigates small-scale farming practices, particularly vannamei shrimp cultivation, in Kudu Hamlet, focusing on how traditional farmers have adopted technology over two decades. Conducted over two years, the study sampled 17 farmers out of 38, analyzing various variables such as age, education, and economic factors. Farmers, predominantly older, have engaged in shrimp farming for nearly two decades, despite fluctuating incomes. Education levels vary, impacting farmers' understanding of new technologies and market trends. Most farmers own their pond land, managing it carefully for shrimp cultivation. Polyculture systems, combining vannamei shrimp with other species, enhance productivity but face challenges like disease outbreaks and market price fluctuations. The choice of feed, predominantly artificial, significantly affects shrimp yields. Farmers also engage in side jobs to supplement income. Overall, enhancing farmers' capacity through education and sustainable practices is crucial for the long-term viability of shrimp farming.
.

Keywords: Vannamei Shrimp; Traditional Farmers; Traditional Ponds, Sustainable Aquaculture Farming

Received 2024-03-09, Accepted 2024-04-29, Available Online by 2024-04-30)

1. Introduction

Indonesia is a significant producer of aquaculture commodities globally, contributing to food security in the aquatic sector. Indonesia ranks second in Asia in terms of aquaculture biota production, with a total output of 14,845 thousand tons or 13.22% of the global production, including finfish (46.9%), mollusks (14%), and crustaceans (9.5%) [1], [2], [3]. Within the crustacean category, Indonesia's aquaculture sector contributes significantly to the export of *Litopenaeus vannamei* or whiteleg shrimp, which remains high annually. Shrimp cultivation in Indonesia contributes 7% of global shrimp production [4].

If traditional farmers are aware that Indonesia significantly contributes to global aquaculture biota production [5], They will be better able to address global environmental issues such as climate change and declining water quality that can impact the sustainability of fisheries cultivation [6], [7]. Traditional fish farmers understanding that Indonesia is a major aquaculture producer globally will help them grasp

the importance of the global market in determining demand and prices for aquaculture products. This will assist in planning production and marketing more effectively [8].

Initially, *Litopenaeus vannamei* shrimp were introduced from Latin America to Indonesia by the government in the early 2000s to support large and medium-scale shrimp industries, which had declined due to the White Spot Syndrome Virus (WSSV) outbreak in black tiger shrimp (Penaeus monodon) [9]. Over time, small-scale shrimp farmers adopted this technology, utilizing traditional ponds for vannamei shrimp cultivation. Generally, vannamei shrimp are more profitable when cultivated on a large or medium scale, either intensively or semi-intensively. However, traditional systems are more beneficial for black tiger shrimp (Penaeus monodon) cultivation [10], [11]. Traditional vannamei shrimp cultivation may be less profitable, but many conventional farmers persist, employing various methods to cultivate vannamei shrimp in their own or leased ponds.

Traditional farmers typically contribute about 5-5% of annual shrimp production, while medium-scale operations contribute 15%. Most large-scale vannamei shrimp production is concentrated among only three to four entities or corporations, accounting for most annual shrimp production (70-80%) [12]. The small contribution from household-scale farmers contrasts with the number of Production Household Units (PHUs) in the majority group. They generally use traditional or extensive technology in earthen ponds with low stocking densities and without precise management practices.

A similar phenomenon occurs in Gresik and Lamongan Regencies, where more than 90% of the current pond area is managed by traditional farmers cultivating vannamei shrimp for decades. Small-scale farmers can participate in shrimp industrialization, as some have good access to certified seedlings and have been using artificial feed [9]. Access to shrimp seedlings from hatcheries and formulated feed is readily available. Based on field observations, the location factor and the presence of numerous intermediary traders for both seedlings and artificial feed aid traditional farmers in obtaining these vital inputs.

As a country with the most significant number of small-scale aquaculture producers globally, alongside Bangladesh and Vietnam [12], scientific studies on small-scale cultivation would greatly benefit its development. Negative stigmas regarding small-scale farming practices, related to environmental unfriendliness and sustainability threats, must be addressed through science, technology, and innovation (STI) initiatives. This research aims to provide a general profile of traditional farmers in Kudu Hamlet engaged in vannamei shrimp cultivation and describe how these traditional farmers' strategies in Kudu Hamlet, serving as small-scale producers, have successfully adopted vannamei shrimp cultivation technology for almost two decades, enabling their ponds to continue producing shrimp.

2. Methods

2.1. Time and Location

This research was conducted over two years, from March 2022 to June 2022 and June 2023 to September 2023. Data was collected in Kudu Hamlet, Weduni Village (Figure 1), Lamongan Regency, East Java.

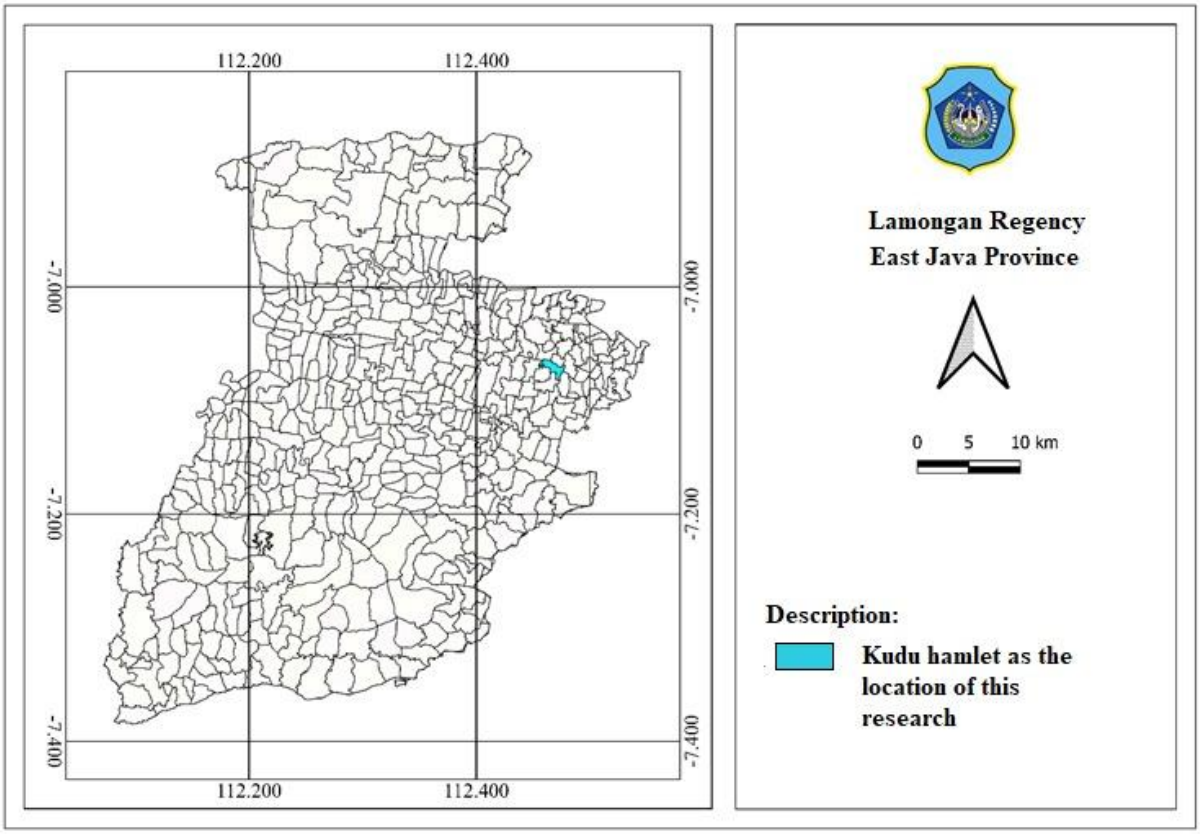


Figure 1. Map of Lamongan Regency

2.2. Population and Sample

The population of this study is traditional farmers from Kudu Hamlet who will be interviewed, with a total of 38 farmers in Kudu Hamlet. Out of the 38 conventional farmers, a sample of 17 individuals was used—the method of selecting the 17 individuals involved matching data from 2021 and 2022. To obtain the sample, the researcher first visited representatives of the Kudu Hamlet management to seek guidance on which residents of Kudu Hamlet were engaged in farming activities. After receiving advice on the names of farmers in Kudu Hamlet, the next step was to visit these farmers and inquire whether they were available for interviews. The criteria for selecting the 17 samples had the same criteria, namely the use of the same number of seedlings and cultivation locations top of Form.

Kudu Hamlet is a hamlet in Weduni Village, part of the Deket District, Lamongan Regency, East Java. Kudu Hamlet covers an area of 410,000 m² and has a population of 125 households, with residents working as farmers, traditional fishermen, and entrepreneurs. There are 88 conventional farmers in Kudu Hamlet. The head of Kudu Hamlet is Mr. Nasim Hadi Prayitno. Facilities in Kudu Hamlet include a practicing midwife, an elementary Islamic school (Madrasah Ibtidaiyah), a Quranic school (TPQ), and the Darul Hikmah Islamic boarding school.

Most farmers in Kudu Hamlet use traditional methods to cultivate vannamei shrimp in earthen ponds, with an average pond area ranging from 5,500 m² to 6,200 m² [13]. These farmers utilize water sources from rivers or rainwater catchment. Characteristics of traditional ponds include the use of simple technology, irregular pond shapes, shrimp stocking densities ranging around 60 per hectare, and the use of natural feeds [14].

2.3. Research Procedure

This study selected main shrimp farmers as respondents, not pond caretakers. The interview sessions lasted an average duration of 60 – 90 minutes. During the interviews with traditional farmers in Kudu

Hamlet, the discussion focused on three variable categories, namely farmer variables, input usage variables, and economic variables in vannamei shrimp cultivation [15]. Farmer variables included the farmer's name, age, education level, side job, address, and the year they started shrimp farming. Input usage variables included probiotics, shrimp seed stocking density, other commodities, feed consumption, and the feed used. Economic variables in vannamei shrimp cultivation included the price of shrimp seeds, feed, probiotics, shrimp harvest quantity, and other commodities, total income from commodities other than vannamei shrimp, and the price of vannamei shrimp per kilogram.

2.4. Research Variables

Table 1. Defined Variables

Variables	Variables Questions During Interview
Farmer Variables	Farmer's name, farmer's age, farmer's educational level, farmer's side job, farmer's address, year of starting vannamei shrimp farming, ownership status of pond land, pond land area, intention to continue farming activities continuously, source of farming knowledge, participation in farmer communities, and the cultivation system used for vannamei shrimp.
Input Usage Variables	the number of shrimp seed stockings, other cultivated commodities, brand or type of feed used, amount of feed used, and use of probiotics.
Economic Variables in Vannamei Shrimp Cultivation	Price of shrimp seeds used, price of feed used, price of probiotics used, quantity of shrimp and other commodities harvested, total income from harvesting commodities other than vannamei shrimp, and price of vannamei shrimp per kilogram.

The categorization of the variable categories above is to facilitate and understand the overall picture of farmers' social and economic profiles, indicating farmers' economic flexibility for the long term, and for the key evaluation of economic efficiency in sustainable shrimp farming.

2.5. Data Analysis

The analysis method used in this study is a descriptive qualitative and quantitative analysis using SPSS and Microsoft Excel tools [16]. Qualitative data to be analyzed using Microsoft Excel will be presented as diagrams, including land ownership status, pond productivity, and side jobs. Meanwhile, quantitative data to be analyzed using SPSS will be presented in t-tests with a significance level of 95%.

3. Results and Discussion

3.1. Profile of Farmers in Weduni Village

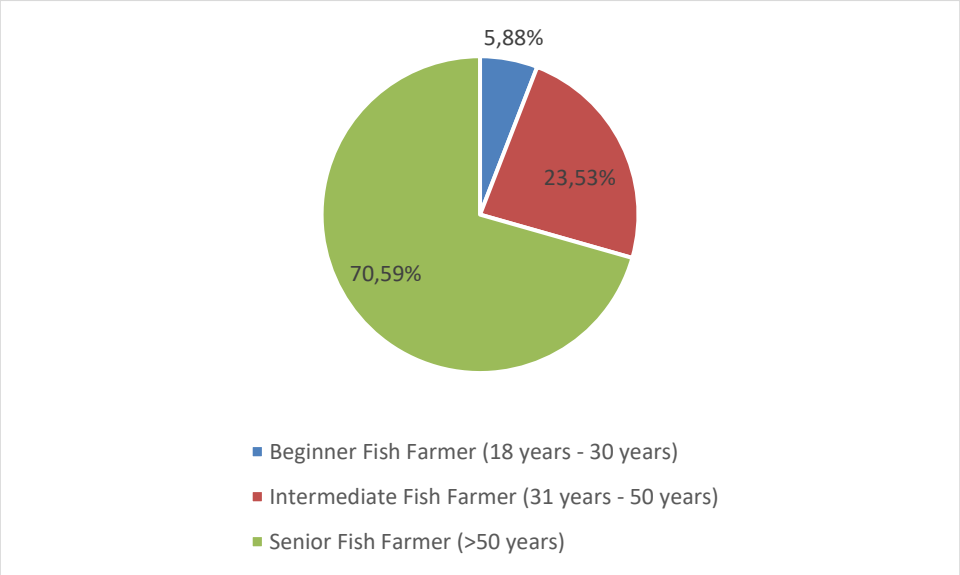


Figure 2. Age of Farmers in Kudu Hamlet, Weduni Village

Farmers' ages can generally be classified into three categories: 18 - 30 years (novice farmers), 31 - 50 years (intermediate farmers), and >51 years (senior farmers) (Sarwana et al., 2019). In Kudu Hamlet, Weduni Village, in 2022, research results regarding the age of farmers yielded an average of 54.29±10.60 years. Farmers in Kudu Hamlet have been engaging in traditional vannamei shrimp farming since 2006, approximately 18 years ago. These farmers engage in traditional farming activities for various reasons: some see it as a flexible additional activity without a fixed schedule, while others make it their primary job due to a lack of alternative options. Despite fluctuating and relatively modest incomes from farming, they remain loyal to these traditional farming practices.

According to previous research, productive farming activities typically occur between the ages of 31 and 50 because farmers are still energetic enough to carry out various farming activities such as preparing land, managing pond operations, and harvesting. Furthermore, previous research indicates that farmers' age reflects their experience in farming; the older the farmer, the more experienced they are in aquaculture [17].

Older farmers typically have years of farming experience. They face various challenges and find solutions to the issues they encounter. By discussing their experiences with younger farmers and the next generation of farmers, this valuable knowledge is passed on, enabling more effective and efficient farming practices. The higher the average age, the more years of experience farmers have in shrimp farming. This is a valuable asset as they have faced various challenges and have extensive knowledge of farming practices. However, without proper succession planning, this knowledge and experience can be lost.

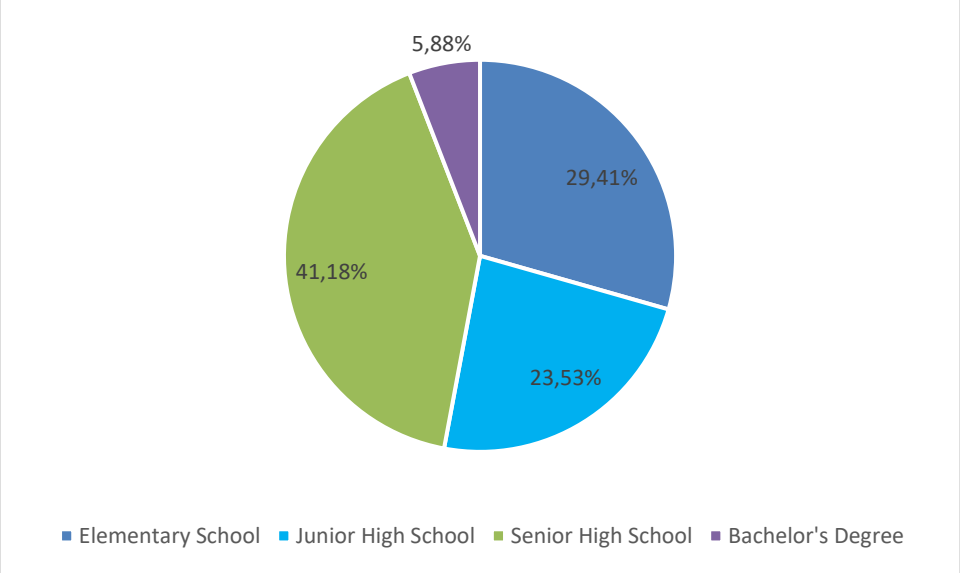


Figure 3. Education Level of Farmers from Kudu Hamlet, Weduni Village, Lamongan Regency

Education is one of the most influential aspects for every individual, as it provides valuable information, insights, and experiences for their future. Education can enhance the quality of individuals' knowledge and social skills. The research diagram in Kudu Hamlet, Weduni Village, depicts four categories of education levels: Elementary School (29.41%), Junior High School (23.53%), Senior High School (41.18%), and Bachelor's Degree (5.88%).

Farmers with higher levels of education tend to have a better understanding of agricultural concepts, resource management, and new technologies. They may be able to assess the benefits and risks of new technologies more effectively. Higher education also enhances farmers' analytical skills to identify problems, understand market trends, and plan more effective agricultural strategies. Higher education often stimulates innovation and creativity. Educated farmers are more likely to try new approaches in farming practices and may be more open to innovative ideas that enhance productivity and sustainability [18].

Table 2. The characteristics of the pond land owned by the farmers of Dusun Kudu, Lamongan.

Aspect	Years 2021	Years 2022
Land Area (m^2)	6.236±4.302,37	5.529±2.826,83
Land Owned By The Farmer Status:		
Owned (%)	88%	88%
Rent (%)	12%	12%

In Table 2, it is explained that 88% of farmers stated that they own their pond land privately, while 12% of farmers stated that they still rent land from others, with rental rates ranging from Rp10,000,000.00 to Rp15,000,000.00 per year. The average land area of Kudu Hamlet farmers is 5,529±2826.83 m². "own pond land" means that farmers own the land they manage. This means they have full control over the management of the pond, including decisions regarding cultivation techniques, investments, and marketing strategies. The ownership of these ponds is typically inherited from their parents and is continuously utilized for vannamei shrimp cultivation.

The land area operated for polyculture shrimp farming by the farmers of Dusun Kudu ranges from 5,529 to 6,236m2 (Table 2). The farmers slightly reduced the area of ponds operated in 2022, possibly due to some land being sold or leased to other farmers, possibly to raise funds for purchasing seed and buying some inputs for farming activities. The ownership status of the land is mostly self-owned (88%), with only a small portion (12%) being leased. Ownership status is related to farming behavior, with landowners generally treating their ponds more carefully than tenants, in terms of the use of medicines and inputs that leave negative residues on soil fertility.

In efforts to enhance the role of traditional farmers in national shrimp production, increasing capacity plays a crucial role. The intended capacity enhancement involves providing traditional farmers greater capacity and production efficiency. They are more inclined to adopt new technologies, better cultivation techniques, and sustainable management practices.

3.2. The Strategy to Persist in Cultivating Vannamei Shrimp Using Traditional Shrimp Farming Systems

Implementing Polyculture Cultivation System

Based on this research, all farmers in Kudu Hamlet apply a polyculture cultivation system to cultivate vannamei shrimp in their ponds. These species include milkfish and tilapia. Polyculture is a cultivation technique in which various biota are raised within one area. By using this method, benefits such as high productivity levels are obtained. Technically, polyculture ponds can be established in almost all areas with sufficient brackish water supply. However, economically, careful consideration of the costs of building and operating polyculture ponds is needed to ensure profitability and avoid losses [17], [19]. The combination of polyculture between vannamei shrimp (Litopenaues vannamei), milkfish (Chanos chanos), and tilapia (Oreochromis niloticus) is considered beneficial because it not only produces three different types of commodities but also because vannamei shrimp acts as a filter to clean the water for the tilapia. Conversely, milkfish and tilapia can act as a balance for the vannamei shrimp population [20], [21].

Furthermore, the long-term sustainability of multi-crop planting systems can be threatened if disease outbreaks occur frequently and cause significant losses. High disease prevalence can jeopardize the economic viability of multi-species cultivation by reducing profitability and increasing operational risks. Therefore, economic analysis of multi-species cultivation should consider the potential impact of disease spread and incorporate risk management strategies to ensure system resilience and sustainability [22].

The next challenge is the limited availability of fertilizers, animal feed, and seeds, which is caused by the frequently changing climate. The prices of feed and fertilizers also pose a challenge to traditional farmers because each year the prices of fertilizers and feed increase, and there are restrictions on their purchase. These challenges can affect the selling prices and the level of productivity of the cultivated commodities.

Table 3. Harvest results of vannamei shrimp (kg/year) cultivated in polyculture by farmers from Dusun Kudu in 2021 and 2022.

Aspect	Years 2021			Years 2022		
	Rerata ± SD	Min	Max	Rerata ± SD	Min	Max
Harvest results of vannamei shrimp (kg/year)	181,18 ± 276,95	35	1200	180,12 ± 348,92	16	1500

Harvest results of vannamei shrimp (kg/year) 181.18 ± 276.95 35 1200 180.12 ± 348.92 16 1500 The average harvest of vannamei shrimp per farmer per year in 2021 and 2022 remains relatively stable at around 180-181 kg/year within an area of approximately 0.5 hectares (as per land area data in Table 2). Over the two-year period, they only carried out one production cycle per year due to constraints posed by the extended seasons and floods longer than in previous years. Generally, the shrimp harvest volume can reflect the level of fertility or productivity of the land in Dusun Kudu, which is considered good, as it is equivalent to the productivity of traditional soil pond land, which generally yields a total vannamei shrimp harvest of 300 kg/year. The factors influencing the productivity of vannamei shrimp in a region are genetic, environmental conditions, and capital in farming activities. The genetic factor refers to whether the cultured vannamei shrimp are resistant to disease attacks or otherwise. Environmental conditions refer to factors that can affect the productivity of vannamei shrimp because if the environment is supportive and soil fertility is maintained, productivity will increase. Capital is one of the factors influencing productivity because it is used to purchase vannamei shrimp seeds, medicines, feed, and fertilizers [23].

A very high variation in total vannamei shrimp harvest was found among farmers. Some farmers obtained very low vannamei shrimp harvests (16 kg/year in 2022), while others achieved high harvest yields (1500 kg/year in 2022). The high variation in total harvest is caused by frequently changing weather conditions and floods that hit the Dusun Kudu area, causing vannamei shrimp to escape to other ponds. The implementation of polyculture farming techniques helps vulnerable farmers who may face crop failures to still earn income. In this study, the number of farmers vulnerable to vannamei shrimp crop failure dominated, accounting for 52% of the total respondents in both 2021 and 2022. These are the farmers who obtained total vannamei shrimp harvests of less than 100 kg/year. Additional income received by farmers from co-cultured species is shown in Figure 4.

Figure 4 shows the production data of co-cultured fish alongside vannamei shrimp.

Figure 4. Harvest Revenue of Milkfish and Tilapia

Kudu Hamlet is one of the hamlets that produce milkfish and tilapia in the Lamongan area. The harvest revenue of milkfish and tilapia in Weduni Village in 2021 and 2022 is as follows: in 2021, the harvest revenue of milkfish was Rp6,741,176±Rp4,870,839.10, and the harvest revenue of tilapia in 2021 was Rp4,705,882±Rp4,815,530.94. Meanwhile, in 2022, the harvest revenue of milkfish was Rp3,920,588±Rp3,905,887.34, and the harvest revenue of tilapia in 2022 was Rp2,423,529±Rp1,518,440.57.

Although the revenue from milkfish and tilapia decreased in 2022, the farmers still earned income from other species, namely vannamei shrimp. The revenue from vannamei shrimp in 2022 reached approximately Rp8,950,000±17,491,102.47. The indicators causing the decrease in milkfish, tilapia, and vannamei shrimp harvests are extreme weather conditions in the Kudu Hamlet area of Weduni Village. Heavy rainfall causes rivers to overflow, leading to many fish and shrimp escaping from several ponds. Some farmers install crickets or nets around the pond walls to reduce the risk of losses due to floods.

External factors influencing income fluctuations include market prices. Market prices are influenced by various factors such as supply and demand, weather conditions, and geopolitical factors. When market prices decrease, the income of farmers and other economic actors also tends to decrease, resulting in income fluctuations. Conversely, when market prices rise, sales increase and can offset minor fluctuations [24]. Additionally, changes in consumer demand can also cause income fluctuations. Changes in consumption trends , customer preferences, or macroeconomic conditions can affect demand for certain products or services. When demand for specific agricultural products increases, farmers' income also increases. However, a decrease in demand leads to negative changes in income [25].

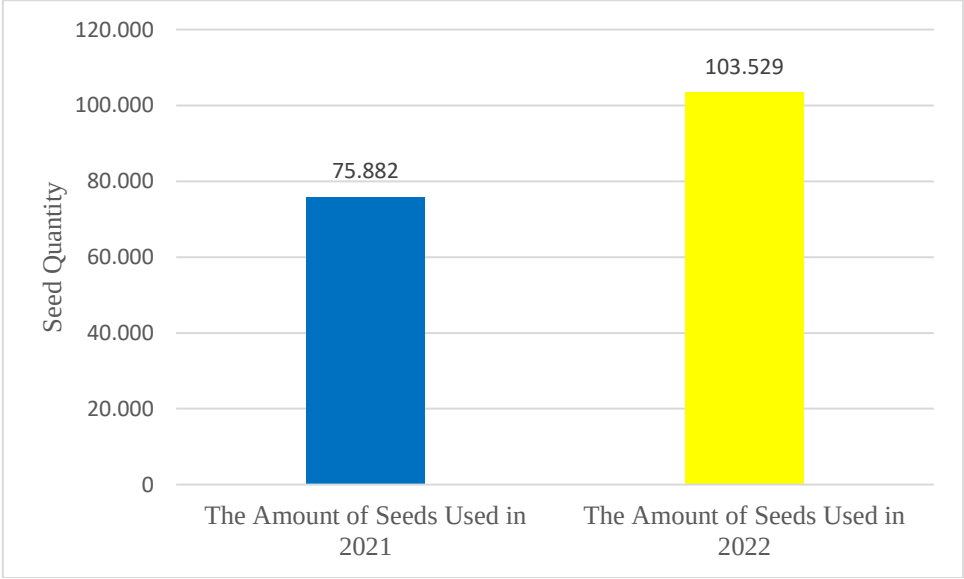


Figure 5. Number of vannamei shrimp seeds used (individuals) in Kudu Hamlet

The above figure illustrates a difference in the number of shrimp seeds used over the two decades. In 2021, the number of seeds used was 75,882±62,405.81 individuals, while in 2022, the number of vannamei shrimp seeds used was 103,529±107,001.24 individuals. The farmers increased the number of vannamei shrimp seed usage in 2022 because they wanted to increase the production of their vannamei shrimp.

Table 4. T-test results of the number of seeds used (individuals) in Kudu Hamlet in 2021 and 2022.

t-Test: Two-Sample Assuming Unequal Variances

	The Amount of Seeds Used in 2021	*The Amount of Seeds Used in 2022*
Mean	75882,35294	103529,4118
Variance	3894485294	11449264706

Observations	17	17
Hypothesized Mean Difference	0	
Df	26	
t Stat	-0,920253844	
P(T<=t) one-tail	0,182950057	
t Critical one-tail	1,70561792	
P(T<=t) two-tail	0,365900114	
t Critical two-tail	2,055529439	

The t-test results indicate that the average number of seeds used in 2022 was 103,529 individuals, while in 2021, it was 75,882 individuals, with a difference between them of 27,647 individuals. The two-tailed p-value obtained is 0.365, indicating that this value is greater than the significance level of 0.05. Thus, the number of seeds used is the same between 2022 and 2021.

Artificial feed is used with various types and brands of feed.

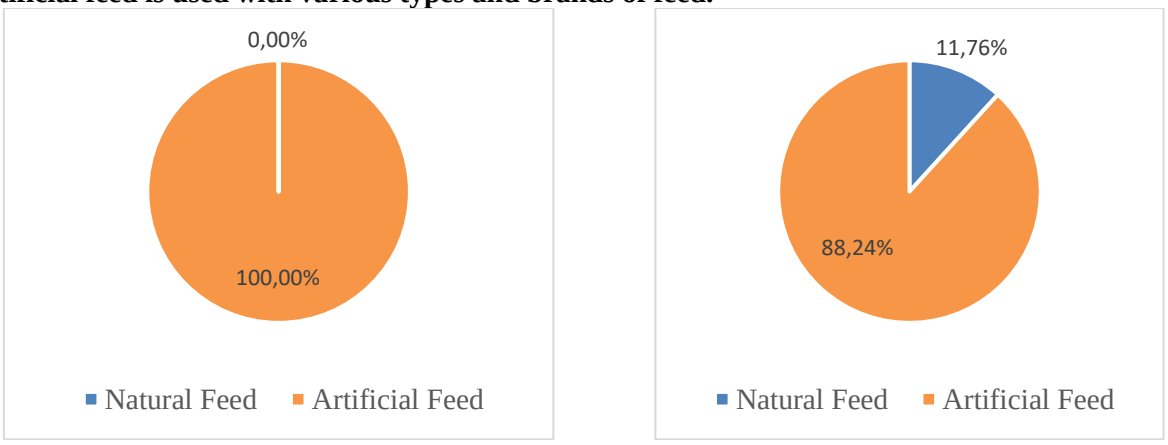


Figure 6. Farmers using feed in 2021 and 2022

Generally, extensive farming tends to rely on natural feed, while semi-intensive or intensive farming uses additional feed to meet nutritional needs optimally. Feed serves as a vital support for the sustenance of life and can enhance the growth rate of the commodities cultivated by farmers [26]. Feeds can be categorized into two types: natural feeds, typically derived from phytoplankton and plankton found in pond waters, and artificial feeds, usually sourced from factories [27].

In Figure 6, it is explained that in 2021, most farmers in Weduni Village used artificial feed. Meanwhile, in 2022, 88.24% of farmers in Weduni Village used artificial feed, while 11.76% used natural feed. Of the 11.76% of farmers using raw feed, it is because they aim to reduce expenses on purchasing feed, as natural feed is readily available from nature and doesn't require additional feed. When comparing the use of natural feed versus artificial feed concerning shrimp yields, the use of natural feed resulted in relatively low yields ranging from 16 kg to 25 kg. In comparison, using artificial feed yielded significantly higher results ranging from 40 kg to 1500 kg.

Stated that food should be distributed evenly so that each shrimp can receive the same amount of food as other shrimp. By feeding evenly, you can avoid competition for food. If conflicts can be avoided, cannibalism can also be avoided. Good shrimp feed requirements include: a) a flat physical surface condition, b) fresh, musty smell, dry and not musty, c) Food packaging is not damaged, d) If pellets are damaged, then pellet stability in water is good. They can survive underwater for at least 2-3 hours. e) Adjust pellet size correctly according to shrimp feeding capacity and shrimp opening size. f) Fresh fish meal has a sweet taste when chewed, and g) appealing. Shrimp feed will be consumed more quickly if using quality feed [28], [29].

Good feed quality is determined by the protein, fat, visible fiber content, and several other nutrients needed for shrimp growth. According In addition to having sufficient protein content, its fat content is also not too high, so that the shrimp get enough feed for their growth [30], [31].

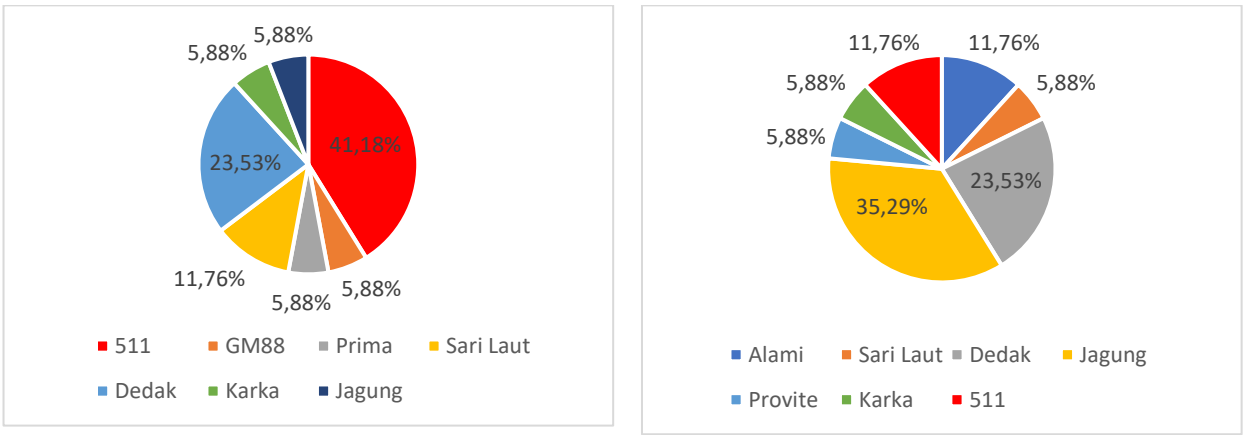


Figure 7. Use of artificial feed brands in 2021 and 2022.

Generally, each artificial feed has its own brand, and each feed has different compositions. In 2021, farmers in Weduni Village used formulated feed brands including 41.18% brand 511, 23.53% bran, 11.76% brand Sari Laut, 5.88% brand GM 88, 5.88% brand Karka, and 5.88% corn. Meanwhile, in 2022, it is known that farmers in Weduni Village used artificial feed brands, including 35.29% corn, 23.53% bran, 11.76% brand Karka, 11.76% natural feed, 5.88% brand Hi-Provite, 5.88% brand Sari Laut, and 5.88% brand 511.

The feed brands used by farmers in Kudu Hamlet, Weduni Village, in 2021 and 2022 can be categorized into two types of artificial feed: animal feed and fish feed. Animal feed consists of bran, corn, brand 511, and GM88. Meanwhile, fish feed consists of Karka, Sari Laut, Prima Feed, and Hi-Provite. Some farmers use animal feed because animal feed is less expensive than fish feed. Thus, farmers can reduce expenses on feed costs. The effect of feeding the cultivated commodities with animal feed is optimal growth and adequate nutrition fulfillment.

Economic Analysis

Table 5. Economic analysis of shrimp harvest (kg/year)

Aspek	Years 2021			Years 2022		
	Rerata ± SD	Min	Max	Rerata ± SD	Min	Max
Shrimp Harvest (kg/tahun)	181,18 ± 276,95	35	1200	180,12 ± 348,92	16	1500
Shrimp Price (Rp/kg)	46.471 ± 6.829,33	30.000	52.000	48.412±6.354,71	45.000	70.000
Total Income From Shrimp Harvest (Rp)	8.652.941 ± 13.896.891,18	1.050.000	60.000.000	8.950.000 ± 17.491.102	720.000	75.000.000

Farmers who obtained the smallest harvest explained that the low harvest was due to In 2021, the farmers obtained shrimp harvests of 181.18 ± 276.95 kg/year with the lowest harvest being 35 kg and the highest

being 1200 kg. In 2022, the total shrimp harvest slightly decreased to 180.12 ± 348.92 kg/year with the lowest harvest being 16 kg and the highest being 1500 kg., ranging from 16-1500kg.

Table 6. Economic Analysis Polyculture Cultivation System

Components	Years	
	2021	2022
Total Revenue	16.923.529,41±9.681.543,877	15.294.117,65±19.956.476,18
Total Expenditure	4.215.529,42±4.922.114,76	4.185.911,76 ± 4.956.611,63
Income	12.708.000 ± 7.758.470,77	11.108.205,88 ± 15.373.833,11

The total revenue referred to in table 3 is the average gross profit from the farming activities conducted in 2021 and 2022, amounting to 16,923,529.41 ± 9,681,543.877 and 15,294,117.65 ± 15,294,117.65 respectively. Total expenditure is the overall expenditure on farming activities such as seed purchases, fertilizer purchases, feed purchases, and probiotic purchases, with an average total expenditure in 2021 and 2022 of 4,215,529.42 ± 4,922,114.76 and 4,185,911.76 ± 4,956,611.63 respectively. Income itself is the net income received by farmers from gross income minus expenditure during the farming activities, with average income in 2021 and 2022 of 12,708,000 ± 7,758,470.77 and 11,108,205.88 ± 15,373,833.11 respectively.

Having a Side Job

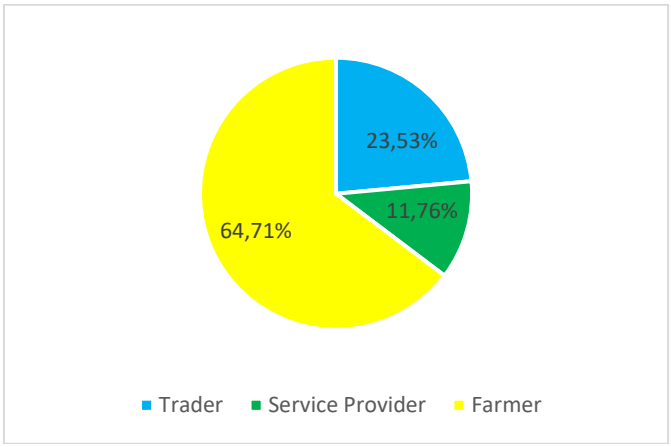


Figure 9. Side Jobs of Kudu Village Fish Farmers

Picture 9 illustrates three categories of side jobs among Weduni Village fish farmers: traders, service providers, and farmers. Fish farmers working as traders account for 23.53%, service providers for 11.76%, and farmers for 64.71%. These fish farmers have side jobs because the income from traditional fish farming activities cannot meet their daily needs, as it often fluctuates from year to year. Therefore, the fish farmers in Weduni Village do not rely solely on fish farming as their profession. Based on the research findings, they stay in the fish farming profession because extensive (traditional) fish farming activities can be done simultaneously with other side jobs.

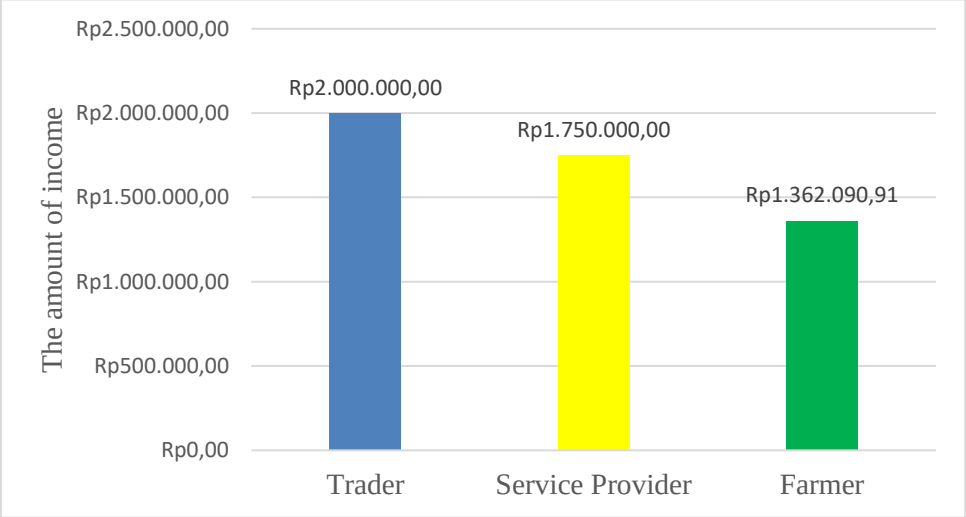


Figure 10. Side Job Income

Side jobs as traders have an income of Rp 2,000,000 ± 1,224,744.87 per month, side jobs as service providers have an income of Rp 1,750,000 ± 353,553.39 per month, and side jobs as farmers have an income of Rp 1.362.090,91±1.111.316,20per month.

Conclusions

In conclusion, the age distribution of farmers in Kudu Hamlet, Weduni Village, reveals a predominantly older demographic, with an average age of 54.29±10.60 years. These farmers have been practicing traditional vannamei shrimp farming for approximately 18 years, with the majority engaging in it as their primary occupation. Despite facing fluctuating incomes, they remain loyal to these traditional practices. Previous research suggests that farmers aged between 31 and 50 are the most productive, leveraging their experience and energy for various farming activities. Moreover, older farmers typically possess extensive experience and knowledge, which they pass on to younger generations, fostering more effective farming practices.

Education plays a pivotal role in enhancing farmers' understanding of agricultural concepts, resource management, and new technologies. Higher levels of education equip farmers with analytical skills to identify problems, understand market trends, and innovate in their farming practices, ultimately improving productivity and sustainability. However, challenges such as limited access to fertilizers, animal feed, and seeds, exacerbated by changing climate conditions, pose obstacles to traditional farmers.

Polyculture cultivation systems, widely adopted by farmers in Kudu Hamlet, facilitate increased productivity through the cultivation of multiple species, including vannamei shrimp, milkfish, and tilapia. While offering economic benefits, this approach also presents challenges, such as disease outbreaks and fluctuations in market prices. The implementation of polyculture systems requires careful consideration of costs and risk management strategies to ensure profitability and sustainability.

The choice of feed type significantly impacts shrimp yields, with artificial feed resulting in higher yields compared to natural feed. Additionally, the brand and composition of feed used by farmers vary, influenced by factors such as cost and nutritional value. Farmers also engage in side jobs to supplement their income, highlighting the need for diversified livelihood strategies.

Overall, enhancing the capacity of traditional farmers through education, technology adoption, and sustainable practices is crucial for improving their livelihoods and ensuring the long-term viability of shrimp farming in the region.

Acknowledgments

We thank Dr Guenwoo Lee and Dr Ayu Pratiwi for the support from Japan Society for the Promotion of Science (JSPS) KAKENHI and Maj and Tor Nessling Foundation are gratefully acknowledged.

Reference

[1] FAO, *The State of World Fisheries and Aquaculture (SOFIA), FAO: Rome, 2022*. 2022.

[2] Z. U. Azizi, F. Farikhah, and A. Aminin, "FAKTOR-FAKTOR YANG MEMPENGARUHI HASIL PANEN UDANG VANAME (Litopenaeus vannamei) DI PERTAMBAKAN KECAMATAN DEKET DAN KECAMATAN KARANGBINANGUN KABUPATEN LAMONGAN," *J. Perikan. Pantura*, vol. 5, no. 2, 2022, doi: 10.30587/jpp.v5i2.3849.

[3] M. A. Syaifullah, F. Farikhah, and A. R. Rahim, "ANALISIS PERBANDINGAN PRODUKTIVITAS KEGIATAN PETAMBAK UDANG VANAMEI (Litopenaeus vannamei) SISTEM TRADISIONAL PADA DUA DESA DI KAWASAN EKOSISTEM ESENSIAL MANGROVE UJUNGPANGKAH," *J. Perikan. Pantura*, vol. 5, no. 1, p. 133, 2022, doi: 10.30587/jpp.v5i1.3846.

[4] F. and A. Organization, "Asian shrimp production fell short of earlier forecast," *Globefish*, 2018.

[5] M. Tajuddin, Ihsan, and Asmidar, "Study Of Design And Composition Of Catch ResultsTrap Netty Tools For Sigeri District Water Pangkep District," *J. Indones. Trop. Fish. ISSN 2655*, vol. 2, no. 1, pp. 86–99, 2019.

[6] A. Saman, E. S. Luhur, S. H. Suryawati, and F. Y. Arthatiani, "Model Permintaan Ekspor Udang Segar Indonesia oleh Pasar Jepang, Amerika Serikat, dan Uni Eropa," *J. Penyul. Perikan. dan Kelaut.*, vol. 15, no. 2, 2021, doi: 10.33378/jppik.v15i2.220.

[7] BPS, "Data Operasional BPS Tahun 2019," 2019.

[8] FAO, "The State of World Fisheries and Aquaculture 2020. Sustainability in action.," *FAO*, 2020, doi: https://doi.org/10.4060/ca9229en.

[9] D. Yi, T. Reardon, and R. Stringer, "Shrimp aquaculture technology change in Indonesia: Are small farmers included?," *Aquaculture*, vol. 493, 2018, doi: 10.1016/j.aquaculture.2016.11.003.

[10] M. Asmild, V. Hukom, R. Nielsen, and M. Nielsen, "Is economies of scale driving the development in shrimp farming from Penaeus monodon to Litopenaeus vannamei? The case of Indonesia," *Aquaculture*, vol. 579, 2024, doi: 10.1016/j.aquaculture.2023.740178.

[11] V. Hukom, R. Nielsen, M. Asmild, and M. Nielsen, "Do Aquaculture Farmers Have an Incentive to Maintain Good Water Quality? The Case of Small-Scale Shrimp Farming in Indonesia," *Ecol. Econ.*, vol. 176, 2020, doi: 10.1016/j.ecolecon.2020.106717.

[12] FAO and WorldFish, *Aquaculture big numbers*. 2016.

[13] Y. P. Hastuti, I. Rusmana, and T. Widiyanto, "Profil tambak tradisional: tekstur tanah, total n-anorganik dan bakteri penghasilnya Profiles of traditional farms: soil texture, total inorganic N and bacteria-producing estate," 2010.

[14] K. Amri, *BUDIDAYA UDANG WINDU SECARA INTENSIF*, vol. 10, no. 5. 2003.

[15] A. S. Nifdhol, Aminin, and Faikhah, "Analisis Produktivitas Tambak UdangVvaname (Litopenaeus vannamei) Semi Intensif di Tiga Provinsi di Pulau Jawa," *J. Technofish*, vol. 7, no. 1, 2023.

[16] D. Firmansyah and Dede, "Teknik Pengambilan Sampel Umum dalam Metodologi Penelitian: Literature Review," *J. Ilm. Pendidik. Holistik*, vol. 1, no. 2, 2022, doi: 10.55927/jiph.v1i2.937.

[17] S. Sarwana, Y. Yumriani, and L. Ismail, "Analisis Budidaya Petani Tambak Terhadap Kondisi Sosial Ekonomi Di Desa Bulu Cindea Kabupaten Pangkajene dan Kepulauan," *Equilib. J. Pendidik.*, vol. 7, no. 2, 2019, doi: 10.26618/equilibrium.v7i2.2683.

[18] S. Ruzzante, R. Labarta, and A. Bilton, "Adoption of agricultural technology in the developing world: A meta-analysis of the empirical literature," *World Dev.*, vol. 146, 2021, doi: 10.1016/j.worlddev.2021.105599.

[19] Y. L. Dhewantara, A. Nainggolan, and I. Amatullah Nabilah, "Sistem Budidaya Polikultur dan Analisis Pendapatan Hasil Budidaya Ikan Bandeng (Chanos chanos) dan Udang Vaname (Litopenaeus vannamei) Secara Tradisional di Desa Hurip Jaya Babelan Kabupaten Bekasi," *J. Ilm. Satya Minabahari*, vol. 8, no. 1, 2022, doi: 10.53676/jism.v8i1.160.

[20] S. Yuliani, "TEKNIK POLIKULTUR UDANG VANAME (Litopenaeus vannamei) DAN

IKAN NILA (Oreochromis niloticus) DI INSTALASI BUDIDAYA AIR PAYAU, KECAMATAN DEKET LAMONGAN," *Ir - Perpust. Univ. Airlangga*, 2015.
[21] A. A. Syahid, M. Ubhan, *Budi daya Udang Organik Secara Polikultur*. Swadaya: Jakarta, 2006.
[22] K. M. Smith, C. C. Machalaba, R. Seifman, Y. Feferholtz, and W. B. Karesh, "Infectious disease and economics: The case for considering multi-sectoral impacts," *One Health*, vol. 7. 2019. doi: 10.1016/j.onehlt.2018.100080.
[23] A. I. Prastianti, "Faktor-Faktor yang Mempengaruhi Produksi Udang Vannamei (Litopeneaus Vannamei) di Desa Pantai Bahagia, Kecamatan Muara Gembong," pp. 49–55, 2021.
[24] U. Kashif *et al.*, "Do agricultural commodity prices asymmetrically affect the performance of value-added agriculture? Evidence from Pakistan using a NARDL model," *Humanit. Soc. Sci. Commun.*, vol. 10, no. 1, 2023, doi: 10.1057/s41599-023-01888-4.
[25] E. H. Brækkan, "Why do Prices Change? An Analysis of Supply and Demand Shifts and Price Impacts in the Farmed Salmon Market," *Res. Gate*, no. July, 2014.
[26] R. Rahman, L. Lahming, and R. Fadilah, "EVALUASI KOMPONEN GIZI PADA PAKAN UDANG FERMENTASI," *J. Pendidik. Teknol. Pertan.*, vol. 4, no. 2, 2018, doi: 10.26858/jptp.v4i2.6617.
[27] D. P. Renitasari, Y. Yunarty, and S. A. Saridu, "PEMBERIAN PAKAN PADA BUDIDAYA UDANG VANAME (Litopenaeus vannamei) INTENSIF DENGAN SISTEM INDEX," *J. Salamata*, vol. 3, no. 1, p. 20, 2021, doi: 10.15578/salamata.v3i1.11259.
[28] K. Darwantin and R. Sidik, "Efisiensi Penggunaan Imunostimulan dalam Pakan Terhadap Laju Pertumbuhan, Respon Imun dan Kelulushidupan Udang Vannamei (Litopenaeus vannamei)," *J. Biosains Pascasarj.*, vol. 18, no. 2, 2016, doi: 10.20473/jbp.v18i2.2016.123-139.
[29] S. Haliman, RW dan D. Adijaya, *Udang Vanname*. Penebar Swadaya. Jakarta., 2005.
[30] D. Heptarina, D., Suprayudi, M. A., Mokoginta, I., & Yaniharto, "Pengaruh pemberian pakan dengan Kadar Protein berbeda terhadap pertumbuhan yuwana udang putih (Litopenaeus vannamei)," *Pros. Forum Inov. Teknol. Akuakultur (Vol. 6)*, 2010.
[31] N. Ulfa, H. Ridar, and Kusai, "Perceptions of fish farmers on striped wallago catfish (Wallago sp.) farming business in floating net cages in Buluh Cina village, Kampar district, Riau province," *Coast. Socio-Economic J.* , vol. 1, no. 3, pp. 49–57, 2020.

Advance Sustainable Science, Engineering and Technology (ASSET)
Vol. 6, No.2, April 2024, pp. 02402024-01 ~ 02402024-08
ISSN: 2715-4211 DOI: https://doi.org/10.26877/asset.v6i2.18284

Design of a Mobile-Based Wedding Information and Booking System using Backend as a Services (BaaS) on android platform

Rafif Mahirssa*, Muhammad Fachrie

Faculty of Sains & Technology, Universitas Teknologi Yogyakarta, Jl. Siliwangi Jl. Ring Road Utara Daerah Istimewa Yogyakarta 55285, Yogyakarta, Indonesia

*rafif.5200411308@student.uty.ac.id

Abstract. Nowadays, wedding organizer services have become a necessity in wedding events. However, it is often found that bookings are made manually to the vendor's location which results in impracticality. This research aims to build an application that brings practicality in the field of wedding organizers. Later this application has an online booking function carried out by wedding bookers and can establish communication with vendors in making wedding bookings. The focus of the research is to provide practicality to bookers and vendors. This research uses the Case Study method, interviews and observations conducted at Omah Kebon Resto as a case study. The application was built based on Android with Cloud Baas services on the Firebase platform for database management. The final evaluation was carried out by distributing questionnaires to respondents for the level of satisfaction of using the application built. The results stated that 37.12% of respondents were very positive, 45.41% stated positive, 13.54% stated neutral, and 3.93% stated negative. So, it can be concluded that respondents feel positive if this application can help practicality in getting information and booking wedding organizer vendors.
Keywords: wedding, wedding booking, wedding organizer, mobile app, online booking

(Received 2024-01-28, Accepted 2024-02-17, Available Online by 2024-03-08)

1. Introduction

Wedding organizer is a service that is used to help brides and families in preparing their wedding planning. Some busy brides and families may seek help from wedding organizers to organize their wedding events without having to bother [1]. Wedding organizer services are often sought after by the public because in addition to making wedding party planning easier, wedding organizers are also very helpful for those who want to avoid the hassle of planning a wedding [2].

Interest in using wedding organizer services is increasing along with people's lifestyle and income. According to Kunjana (2020), in 2020, this interest increased by 20-30%, which resulted in intense competition among wedding organizer service providers. Therefore, business people in the wedding organizer industry are expected to have special strategies to survive in this increasingly competitive era [3].

When promoting wedding services, a wedding organizer faces the problem of not being able to present complete wedding package information. Some of the problems that arise include difficulties in

managing the latest wedding package list and processing transaction data manually, which can cause misinformation on order details and difficulty in finding customer order data. In addition, the manual way of recording and calculating also has the potential to cause errors in determining business profits and losses [4].

Applications are the right solution. The role of applications is an important support to help advance business strategies [5]. Making applications can aim to provide practicality and convenience for wedding event bookers who want to get information or place an order [6]. This will also have an impact on employees in managing wedding orders more easily. Furthermore, this application will provide information in the form of menu details provided by the vendor, this provides information to the booker in order to get an overview before booking the event through the application later.

Before conducting research, a literature review has a very important role as a reference for gathering information and aspects needed later. The literature review is also considered important because it is a foundation that explains why researchers choose a particular theme or title. In addition, the literature review serves as a foundation for outlining theories, findings, and materials relevant to the research, so that it becomes a strong basis for conducting the research conducted. Thus, the literature review provides an in-depth and contextual framework for the work to be reported [7].

1. In research conducted by R. Aulianita, entitled "User Center Design in Building a Website-Based Wedding Organizer", the principles of User Centered Design are applied to the website. This research aims to provide a platform used by wedding organizers to conduct promotions, thus allowing prospective brides to make reservations more easily [8].
2. In research conducted by S. A. Haswir and K. Budayawan, entitled "Designing an Information System for Management of Auction Services", this research implements the CodeIgniter framework. The result is an information system for managing pampering services as a promotional and ordering media that allows reservations and payments to be made online via the website [9].
3. In a study conducted by T. Pintar, entitled "Application of Sms Gateway to the Wedding Property Booking System using RESTful Web Service at Puspita Wedding", implementing RESTful Web Service and SMS Gateway. The result is a structure that can be used in various modern programming languages, known as the JSON format. The website made it easier for the admin to get data and record rental transactions efficiently. The implementation of RESTful Web Service and SMS Gateway using Gammu can reduce the SMS Gateway hosting rental budget [10].

From all the explanations that have been done by previous researchers, it can be concluded that the discussion of applications that support businesses in the field of wedding organizers is an interesting topic. Therefore, in this paper, the research conducted is to design an Android-based application with the Flutter framework by utilizing the Baas service on Firebase to store data about the information and menus provided later, the application to be built will be Android-based using the Flutter framework and the Backend as a Services (BaaS) service on Firebase, which aims to facilitate application connection to the database in real time. The use of Firebase services from Google was chosen to provide stability when the application connects data to the database. This is expected to increase user responsiveness in using the application.

This research implements the case study method, which is an in-depth exploration of a "finite system" or "diverse cases" in a specific context. This method involves in-depth data collection from various sources of information to understand and analyze the implementation of the application that has been built. This case study is expected to provide comprehensive insights This research implements the case study method, which is an in-depth exploration of "limited systems" or "diverse cases" in a particular context. This method involves in-depth data collection from various sources of information to understand and analyze the implementation of the application that has been built. This case study is expected to provide comprehensive insights [11].

2. Methods

This stage contains a discussion of the application that was built with a discussion based on the method used, namely Case Study. Case Studies are found in many studies that play a role in assisting in

presenting research guidelines from simple to complex levels. This method has a design with an arrangement of stages tailored to research needs starting from introduction, research objectives, literature study, subject selection, system design, application development, data collection, data analysis, and conclusions.

This research started by finding out about the development of wedding booking industry trends. The aim is to make it easier for wedding bookers to access information and make bookings with wedding vendors efficiently. The next step involved searching and learning from previous research related to the topic to be researched. The case study location was chosen at Omah Kebon Resto, a wedding organizer service provider. Interviews and observations will be conducted to collect the necessary data. After that, based on the results of observations and discussions with the vendor owner, a system that suits the vendor's needs will be designed before entering the application development stage. The system development process is then carried out in accordance with the design and planning that has been made previously. Data is collected after the customer tries the application, and questionnaires are filled out to obtain application trial data. The final stage involves analyzing and assessing the data that has been collected from filling out questionnaires, as well as concluding and discussing the final results of the research that has been carried out.

2.1. Research Framework

At At this stage, researchers conducted observations and interviews. Observations were made by directly observing interactions between employees and customers at Omah Kebon Resto to understand how information is delivered, while interviews were conducted with the owner and employees of the restaurant to gain insight into the menu and other needs that support website development. Thus, researchers can compile a research framework as shown in Figure 1 below:

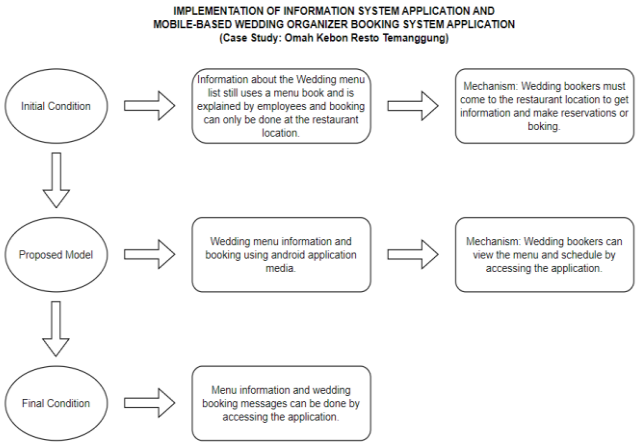


Figure 1. Reseach Framework

2.2. Research Data

The data source is the subject from which the data is obtained. In this study, researchers used two data sources, which consisted of.

1. Primary data

   Primary data is data obtained by means of being collected by researchers directly from the main source. The primary data sources in this study are menu lists, information, price lists.

2. Secondary Data

   Secondary data is data collected by researchers from pre-existing sources of information. Secondary data is generally in the form of graphs, diagrams, and tables. The secondary data sources in this study are various journals as references.

2.3. Application Design

Can be seen in Figure 2 shows the Use Case diagram of the application made to see the service model contained in the system. In the diagram, there are 2 actors, namely employees and customers. Employees have a role in inputting the menu available at the wedding organizer, as well as taking care of the status of orders that will be sent to customers. In this Use Case diagram, there is a scope of the system being developed and serves as a medium for presenting the roles of actors who are outside the box. The role of employees is to be able to add menu categories, add menus, view order schedules, view customer orders, and send confirmations to customers. Customers can play several roles, such as viewing the menu in the application, viewing available information, adding menus to the list, viewing the list in the basket, filling out the order schedule, checking out, and viewing the order confirmation status.

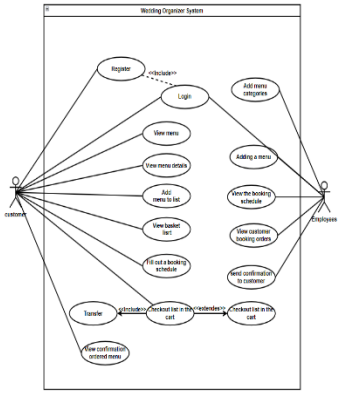


Figure 2. Use Case Diagram

2.4. System Flows

In designing this application using flowchart. Flowchart itself is used to show the overall system workflow. In general, the flow of the wedding organizer booking system can be seen in Figure 3 below.

Figure 3. Flowchart

Customers can access the web or app and login. If they do not have an account, they can register first before logging in. After login, customers will be directed to the first home page. Here, they can access the menu by selecting the desired menu category, and then view the product details and choose to add it to the cart. Customers can view products that have been added to the cart and check the status of orders that have been checked out. They can also ask questions about the wedding menu via chat and customize their profile. Transactions can be made after checkout, and customers will receive a confirmation from the wedding after the transaction is completed. On the other hand, employees need to login before accessing the wedding menu dashboard. They have the authority to add, delete, and edit wedding categories and menus. In addition, employees can view a list of orders that have been placed by customers, as well as confirm transaction status. Interaction with customers can also be done through the chat feature.

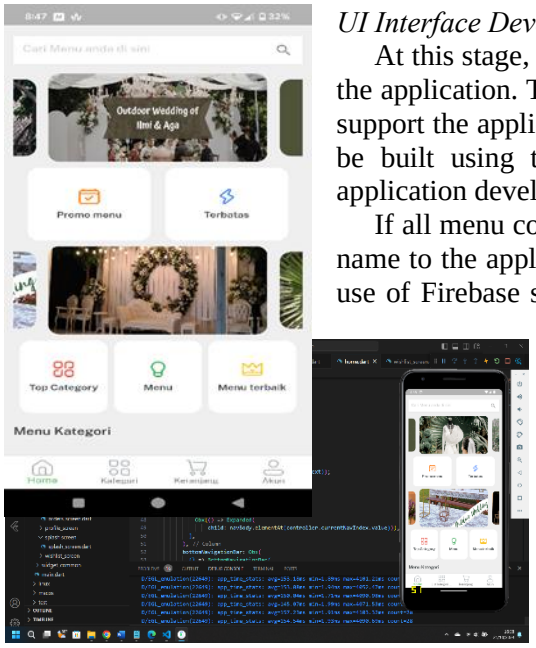


Figure 4. Application Development Process.

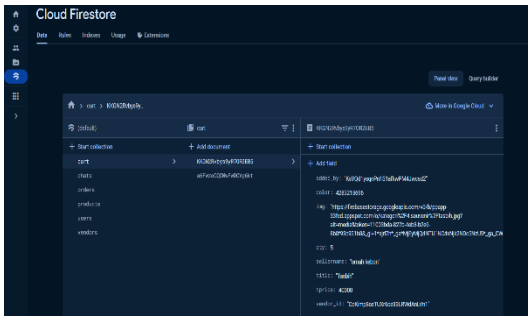


Figure 5. Database usage

UI Interface Development

At this stage, the next activity carried out is to create a menu display in the application. The display that will be made includes the parts needed to support the application when used later. Furthermore, this application will be built using the Flutter framework. Figure 4, and Figure 5 is the application development process using Flutter.

If all menu components have been completed, the next step is to give a name to the application, an explanation of the application details, and the use of Firebase services to manage user authentication in the application later.

The application development stage is carried out using VS code which is a type of text editor with the addition of the Flutter framework extension. Furthermore, the finished application will be built app, so that it can be tested on several Android devices.

3. Results and Discussion

3.1. Result

The results of the application interface can be seen in Figure 6. When the customer/user successfully logs in, they will be able to access various content menus in the application. In Figure 7 the wedding menu category page separates the wedding menu catalog. On the profile menu, several menus can be accessed as in Figure 8.

3.2. Application Testing

At this stage is the final part as in the Case Study method, namely evaluation and discussion section. evaluation stage has the aim of the application that has been made it can find out how feasible the application can carry out its performance. The next stage that carried out is the creation of a questionnaire to determine the level satisfaction with the application built. The following is Table 1

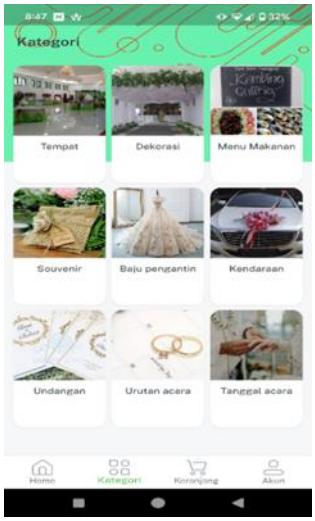


Figure 7. Category Pages

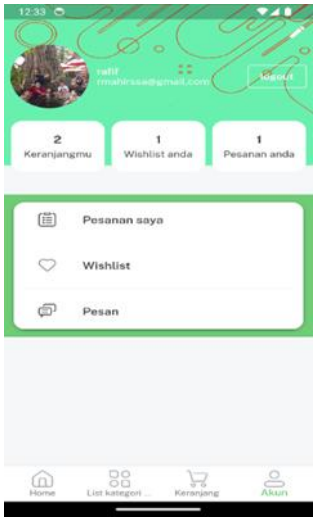


Figure 8. Profile Pages

found the The testing so that will be of user being which

displays the questions from the questionnaire distributed to customers who have tested the application. From the incoming questionnaire, 23 participant votes were collected.

Table 1. Questions regarding user satisfaction level

No	Question
1	What do you think about the display contained in this wedding booking application?
2	Does this app offer convenience in ordering wedding menus?
3	Is the app running properly, with no issues experienced?
4	Does this app help you get wedding menu information more easily?
5	What do you think about the performance response of this application?
6	How easy is it to use this app?
7	Does the app help you to place an order?
8	Are you satisfied with the features this app offers?
9	How easy is it for you to place an order on this app?
10	What do you think about the completeness of the wedding menu provided?

From the distributed questionnaires, the calculation and compilation results are listed in Table 2 below. From a total of 23 responses to 10 questions, there are variables A to D, with values of 4 to 1. The results are as follows: Option A received 85 votes, Option B received 104 votes, Option C received 31 votes, and Option D received 9 votes. After conversion into percentages, 37.12% chose A, 45.41% chose B, 13.54% chose C, and 3.93% chose D. These results show that the majority of users chose option B, signaling a positive response of 45.41%. This indicates that this application can help the efficiency of work in the wedding organizer business. Evaluation of the operation and use of the application showed that the questionnaire fillers did not find any obstacles and considered the application to run well. The information presented and the ease of access to the wedding menu were considered very helpful by the questionnaire fillers. Overall, the application's performance was rated as fast and responsive by the users.

Table 2. User satisfaction questionnaire results.

No pertanyaan	Pilihan jawaban				total
	A	**B**	**C**	**D**	
1	9	10	3	0	23
2	7	9	6	1	23
3	12	8	3	0	23
4	10	10	2	1	23
5	17	6	0	0	23
6	9	12	2	0	23
7	6	10	5	2	23
8	8	11	4	0	23
9	5	17	1	0	23
10	2	11	5	5	23

The results of the questionnaire showed the majority of users chose option B, indicating significant satisfaction with the features of the wedding organizer app. This confirms the importance of research in understanding user needs and the potential of technology to improve industry efficiency. However, the limitations of the questionnaire need to be recognized in the in-depth understanding of user preferences. With the analysis of the questionnaire results, this research provides important insights for technology development in the industry, enabling the design of more effective and relevant solutions to improve the user experience and efficiency of wedding event organization.

4. Conclusion

Based on the discussion and research results that have been carried out in the process of designing an Android mobile-based wedding menu information ordering and accessing application, the use of the Case Study model in its implementation is a very useful guide in exploring the dynamics of using this application. The evaluation revealed that most customers were satisfied with the app, with 37.12% expressing a very positive level of satisfaction and 45.41% expressing a positive level of satisfaction. However, there were also some who felt neutral (13.54%) or negative (3.93%) about the app. Thus, the potential benefits of this application to the efficiency of placing orders and accessing menu information provided by wedding organizer vendors are becoming increasingly clear

This reinforces the broader impact of this research on the wedding organizer industry. The app not only makes it easier for customers to plan weddings, but can also improve the efficiency and quality of services provided by wedding organizer vendors. Stakeholders, including vendors, brides-to-be, and professionals in the industry, can benefit from the app in improving user experience and operational efficiency.

In this context, recommendations for practitioners and policymakers include efforts to continuously improve and update the UI to make the application more interactive and attractive to users. In addition, the addition of a web-based system could expand the application's access range. The addition of features also needs to be considered in accordance with future development needs, so that this application remains relevant and can meet the demands of a growing market. Thus, this research not only contributes directly to the development of the application, but also provides useful insights for the development of the wedding organizer industry as a whole.

Acknowledgements

The author would like to thank Omah Kebon Temanggung for supporting this project.

References

[1] M. Aman and Suroso, “Pengembangan Sistem Informasi Wedding Organizer Menggunakan Pendekatan Sistem Berorientasi Objek Pada CV Pesta,” *Jurnal Janitra Informatika dan Sistem Informasi*, vol. 1, no. 1, pp. 47–60, Apr. 2021, doi: 10.25008/JANITRA.V1I1.119.

[2] R. Salma, N. Samsiar, and I. D. Mayangsari, “Analisis Strategi Komunikasi Pemasaran Megrashy Wedding Planner Dalam Meningkatkan Penjualan Di Era Pandemi Covid-19 Pada Aplikasi Tiktok,” *Management Studies and Entrepreneurship Journal (MSEJ)*, vol. 3, no. 6, pp. 3431–3437, Nov. 2022, doi: 10.37385/MSEJ.V3I6.1136.

[3] I. Aprilia, Z. 1✉, Z. Abidin, M. Rifai, and K. Mardiyah, “Pengaruh Komunikasi Pemasaran Digital Pada Media Sosial Instagram Terhadap Brand Awareness Dan Keputusan Menggunakan Jasa Wedding Organizer Teman Manten,” *Innovative: Journal Of Social Science Research*, vol. 3, no. 4, pp. 638–646, Aug. 2023, doi: 10.31004/INNOVATIVE.V3I4.3569.

[4] A. Ishak and N. Pakaya, “Sistem Informasi Wedding Organizer Berbasis Android,” *Jambura Journal of Informatics*, vol. 3, no. 2, pp. 97–108, Nov. 2021, doi: 10.37905/jji.v3i2.11746.

[5] A. Shobirin, G. Aryotejo, and D. Danang, “Design and Development of Tourism Geographical Information System of Semarang City Based On Android Mobile,” *Advance Sustainable*

Science, Engineering and Technology, vol. 2, no. 1, p. 0200103, May 2020, doi: 10.26877/ASSET.V2I1.6018.

[6] H. Fuad, A. Budiman, and D. Kurniasari, "Perancangan Sistem Informasi Pemesanan Paket Pernikahan Berbasis Web Study Kasus Di Wedding Organizer PJ Management," *JURNAL SISFOTEK GLOBAL*, vol. 8, no. 2, pp. 2088–1762, Sep. 2018, doi: 10.38101/SISFOTEK.V8I2.202.

[7] M. Ridwan, B. Ulum, F. Muhammad, I. Indragiri, and U. Sulthan Thaha Saifuddin Jambi, "Pentingnya Penerapan Literature Review pada Penelitian Ilmiah," *Jurnal Masohi*, vol. 2, no. 1, pp. 42–51, Jul. 2021, doi: 10.36339/JMAS.V2I1.427.

[8] R. Aulianita, "User Center Design Dalam Membangun Wedding Organizer Berbasis Website," *INFORMATION MANAGEMENT FOR EDUCATORS AND PROFESSIONALS: Journal of Information Management*, vol. 4, no. 1, pp. 31–40, Dec. 2019, Accessed: Oct. 20, 2023. [Online]. Available: http://101.255.92.196/index.php/IMBI/article/view/1193

[9] S. A. Haswir and K. Budayawan, "RANCANG BANGUN SISTEM INFORMASI PENGELOLAAN JASA PELAMINAN (Studi Kasus: Gadih Nan Tongga Wedding)," *INCARE, International Journal of Educational Resources*, vol. 2, no. 3, pp. 291–301, Oct. 2021, doi: 10.59689/INCARE.V2I3.287.

[10] R. F.-J. T. Pintar and undefined 2022, "Penerapan SMS Gateway pada Sistem Pemesanan Properti Pernikahan menggunakan RESTful Web Service di Puspita Wedding.," *teknologipintar.org*, vol. 2, no. 12, pp. 2022–2023, 2022, Accessed: Nov. 08, 2023. [Online]. Available: http://teknologipintar.org/index.php/teknologipintar/article/view/311

[11] D. Assyakurrohim, D. Ikhram, R. A. Sirodj, and M. W. Afgani, "Metode Studi Kasus dalam Penelitian Kualitatif," *Jurnal Pendidikan Sains dan Komputer*, vol. 3, no. 01, pp. 1–9, Dec. 2023, doi: 10.47709/JPSK.V3I01.1951.

Advance Sustainable Science, Engineering and Technology (ASSET)
Vol. 6, No.2, April 2024, pp. 02402025-01 ~ 02402025-08
ISSN: 2715-4211 DOI: https://doi.org/10.26877/asset.v6i2.18435

Portable Low-Cost Home Sleep Monitor using Wemos D1 Mini

Pramudita Hanggara[1], Ade Septian Alfianto[2], Louis Putra Purnama[2], Ivan Tanra[1*],

[1]Department of Electrical Engineering, Universitas Kristen Krida Wacana, Jl. Tanjung Duren Raya No.4, RT.12/RW.2, Tj. Duren Utara, Kec. Grogol petamburan, Kota Jakarta Barat, Daerah Khusus Ibukota Jakarta 11470, Indonesia

[2]UKRIDA Science Centre, Universitas Kristen Krida Wacana, Jl. Tanjung Duren Raya No.4, RT.12/RW.2, Tj. Duren Utara, Kec. Grogol petamburan, Kota Jakarta Barat, Daerah Khusus Ibukota Jakarta 11470, Indonesia

*ivan.tanra@ukrida.ac.id

Abstract. Sleeps disorders are a common disease overlooked by many people. Sleep disorder have many types and kinds and often associated with other severe illness such as diabetes, stroke, obesity and many others. Sleep monitors are one of many ways to read all parameters related to sleep and detect sleep disorders the subject has, however access to sleep monitor still expensive and tough to come by. Because of that accessibility, countless development of home sleep monitor occured around the world. However, most of that device still hard to operate and some of them gives error readings of parameters. Based on this, a portable low-cost home sleep monitor was developed using Wemos D1 mini as a microcontroller, MAX30102 as an oxygen level sensor, MPU6050 as an accelerometer, DS18B20 as a breathing flow sensor, and MAX9814 as a microphone. Each of the sensors read and give value to microcontroller and store the data to cloud and display the result in user's gadget. The aim of this development is to detect sleep disorders associated with each reading of sensors used and determine sleep quality as an early detection of symptoms before referring to professional relaled to sleep disorders or doctors.

Keywords: Sleep Monitor, Diagnosis, Portable, Wemos D1 Mini, MAX30102, MPU6050, DS18B20, MAX9814, Sleep Apnea, Willis-Ekbom.

Received 2024-02-13, Accepted 2024-04-24, Available Online by 2024-04-30)

1. Introduction

Sleep less than 7 hours per night often associated with risk of obesity, diabetes, stroke, mental distress and impairs cognitive performance which indirectly increase the chance of many kinds of accidents and loss of productivity caused by hard to focus[1]. The cumulative of sleep loss stretch to physical and mental health problems such as reduced memory function, negative mood states, obesity, hypertension, and reduced immune response[2]. Lacks of sleeps only has negative effect to humans health which result in reduced quality of life, given that facts, total of sleep has spiked down compared to 50 years ago, approximately 20% adult around the world right now has nocturnal sleep disorders and it affects not only the healths but economicly, sleep disorder related treatment and expenditure in the United States costs around US$165 billion per year, in comparison treatment of heart failure, stroke, and asthma costs around US$20 to US$80 billion per year[3].

Sleep disorders are more common and problematic than many thought, many countries and researcher have tried to solve the problem and have come to a realization which is sleep study using sleep monitor. Sleep monitor as its name suggest is monitoring subject in sleep and collect data from all kinds of sensors after a few nights of sleep study at hospital or sleep center and professional will process the data and figure out any kinds of sleep disorders the subject has from all data collected and events happening when subject sleeps. Polysomnography or PSG for short is the current standard for measuring sleep, which measure all kind of parameters such as EEG, eye movements, muscle activity, heart rate, and breathing rate[4] to monitor individuals sleep, which the individuals typically spend the night in sleep laboratory to ensure effective sleeping environment and employed by numerous numbers of surface electrode to measure the parameter under the surveillance of sleep technician[5]. PSG has many restrictions to ensure correct monitoring of subjects which can affects subject's sleep in many ways and hinder the process of sleep study[6]. Sleep study without many restrictions are what many people try to invent to ensure as little as possible effect to subject sleeps. Now many inventions have emerged and sleep monitoring from home is possible with the main objective of this invention is the same as prologue to PSG, detecting early symptoms of common sleep disorder such as sleep apnea[7] and excessive daytime sleepiness[8], from there subject can be treated quickly according to symptoms or consult with professional such as sleep study center or doctor.

Early detection of any sign of sleep disorder such as sleep apnea and willi-ekbom disease are crucial[9] as the scope of this development heavily depend on that point and user friendliness for using the product at home as well as affordable. The significance of proposed device are detecting signs of sleep disorder at early stage without disturbing the sleep of user and more affordable than going to hospital or sleep center for diagnoses. The proposed device give affirmation to user about their sleep habits and the needs to go to hospital or sleep center for professional treatments or not.

2. Methods

2.1. Proposed System

Home sleep monitor need a few parameters values to determine any kinds of sleep disorders, such as breathing flow, movement of body parts, and sounds[10]. Shown in Figure 1, sleep monitoring system designed with sensing system to get the parameter values using accelerometer, oximeter, microphone, and thermometer with 2 microcontrollers divided as 2 unit. One of the units with breathing rate sensor and microphone as sound sensor with a microcontroller is called mask unit, and the other is called hand unit with heart rate sensor and

accelerometer as motion sensor with a microcontroller. Data from all of these sensors will be sent to corresponding microcontroller which microcontroller 1 with accelerometer and heart rate sensor and microcontroller 2 with microphone and thermometer. After data transfer to microcontroller, all data will be sent to database as organized number in tables and displayed as graphics.

Each parameters have a sensors correspondent to each of them. Wemos D1 Mini as microcontrollers, MPU6050 for the accelerometer, MAX30102 for the oximeter, DS18B20 as the thermometer for breathing flow, and MAX9814 as the microphone. Battery 18650 with 3,7 Volt and 3000MaH for the power supply with 18650 chargers to replenish the power.

Wemos D1 Mini is a low-cost and low-power microcontroller that integrated with WiFi and Bluetooth based of ESP8266[11]. Chosen for this project because of its affordability and function.

MPU6050 is an accelerometer sensor which work with low voltage, makes it the suitable sensor to capture the acceleration of a body parts attached to it. The parameters to determine the acceleration of body parts is 3 axis of sensor positon which is X-axis, Y-axis, Z-axis, from the 3 positions change's the sensor calculates it against the time and gives out the acceleration per seconds[12]. Restless legs or hand syndroms are the sleep disorder which can be detected by data captured with MPU6050 as they measure movement of limbs attached to sensor[13].

MAX30102 is a high sensitivity pulse oximeter sensor able to capture many kinds of parameters such as heart rate, oxygen levels, and temperature. The parameters of oxygen level can indicate for hypoventilation sleep disorder which the oxygen levels drop and accompanied with episode of shallow breaths [14].

Microphone by MAX9814 used to capture noises around subject to determine parameters of noises that can affect sleep at certain levels. Noises from subject such as snoring can be captured too and included in sleep disorders and snoring sound can affects the sleeps as well[14].

DS18B20 is a low-cost 1-wire progammable temperature sensor operates within a wide temperature range from -55°C to +125°C with a decent accuracy of ±0.5°C. Communication with the sensor follows the 1-Wire method, and the temperature data is stored in a 2-byte register within the sensor[15].

Computer will be connected to the device through WiFi and read the data result from all the components mentioned above which all connected to Wemos D1 Mini to operate and gather all the parameters. The parameters will be transferred to Google Sheets as database and displayed as graph better for analyze and easier understanding. Google sheets is a free, cloud-based spreadsheet application provided by Google. It enables users to create and edit on spreadsheets seamlessly from any device[16]. It's used for data analysis, visualization, and record-keeping in real time for the proposed device.

2.2. Flowchart

The device utilizes the Arduino IDE as its application platform to program the device to read the parameters and connect to WiFi. The device's flowchart is shown in Fig. 2 and as shown the device will boot up and start initialization of the sensors, if the initialization is successful, all the sensors will start to boot up and will try to initialize again if the earlier initialization failed. The next in flowchart is to connect to network with WiFi and to database, if connected successfully the process will go on and the sensors will start to capture data to microcontroller and will send it to database through the network, if the device failed to connect to WiFi and database then the device need to be restarted and if the same problem keep occurring the code for WiFi SSID and password need to be double-checked. The process still going on and the data from database will be processed and displayed in the computers in the form of graph. The displayed data still need to be manually analyzed as the data don't have significant difference and hard to organize with automated process. Each sensor's data provides distinct indicators of various sleep disorders, with no inter-sensor correlations and calculated to provides user-friendly graph.

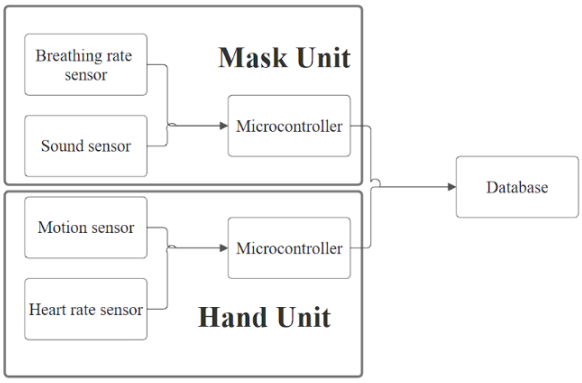


Figure 1. Diagram Blocks of Portable Sleep Monitoring System

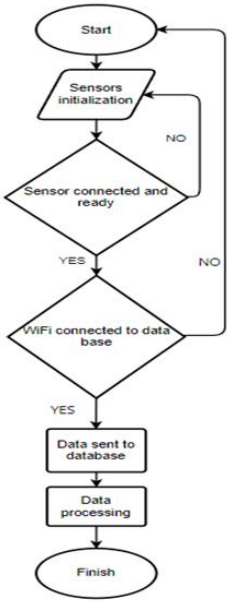


Figure 2. Device's Flowchart

2.3. Schematic

The schematics of the devices is two different unit, shown in Fig. 3, the corresponding sensors is the MPU6050 and MAX30102 and 3 pins of 2 which is battery, switch, and LED. The sensors will be connected to D1(SCL) and D2(SDA) as the communication pin. Battery connected to microcontroller through switch and LED as power indicator for the hand unit.

Mask unit accompanied with microphone and thermometer and the same battery, switch, and LED. As shown in Fig. 4, the microphone connected to A0 to capture analog inputs and thermometer connected to D1(SCL) and D2(SDA) as communication pin. Battery connected to microcontroller through switch and LED as power indicator for the mask unit.

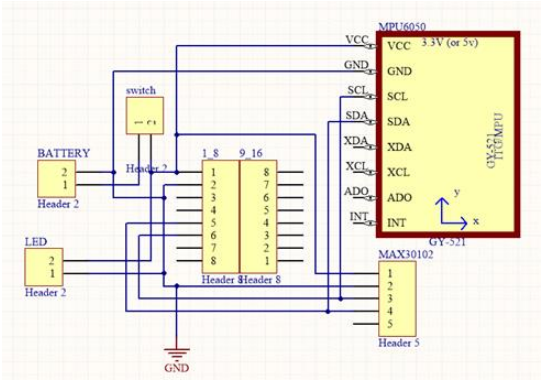


Figure 3. Circuit Schematic Hand Unit.

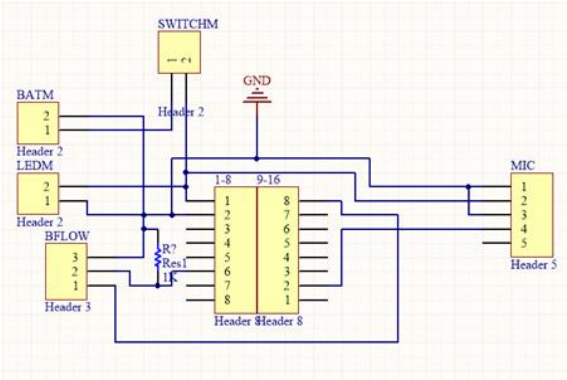


Figure 4. Circuit Schematic Mask Unit.

2.4. Design Implementation

The design of device is separated as 2 unit as shown in Fig. 5. Fig. 5 (a) is hand unit will be attached to around wrist like a watch and Fig. 5 (b) is mask unit in the form of oxygen mask with the device on its bottom and equipped as an oxygen mask should.

(a) **(b)**

Figure 5. (a) Prototype Design Hand Unit, (b) Prototype Design Mask Unit.

2.5. Tool Installation

All unit will be worn according to each unit design as shown in Fig. 5, all unit designed to be easy to use and efficiently gather data from corresponding sensors. Hand unit has a housing of square for the wrist and has a housing for the MAX30102 sensor on the finger. Placement of MAX3012 sensor for reading need specific body part for the optimal readings, making the placement below finger the most suitable[17]. Mask unit has a housing for microcontroller and a mask for the sensors to read and gather data.

3. Results and Discussion

3.1. Oxygen Level Test

The device will be worn exactly shown before and all data from the sensors will be forwarded to microcontroller and will be send to database in the form of oxygen saturation will be send too with the same sensor as heartrate where the oxygen levels normaly stays within 95-100% for a healthy individual and considered dangerous if the reading stays at below 94%[18].

In Fig. 8, the measurement shows sometimes that the reading below 90% can be considered errors according to data only seconds of reading below 90% then back to around 95-100%. In the span of 2-4 hours of reading the sensors, the reading from sensors shows some error but overall stable reading of a healthy individuals. In Fig. 9 shown the validation of data between oxymeter and sensors MAX30102 measuring oxygen level and from 1 minute of data gathering, shows different value of 1-2% in SpO2[19].

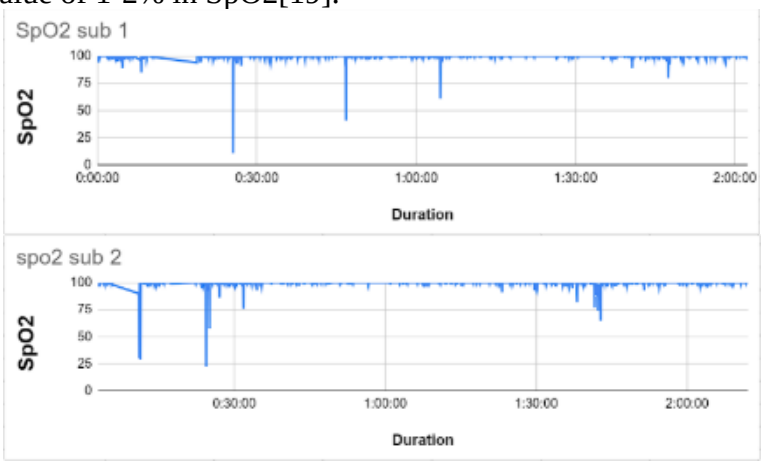


Figure 6. Measurement of Oxygen Levels from MAX30102.

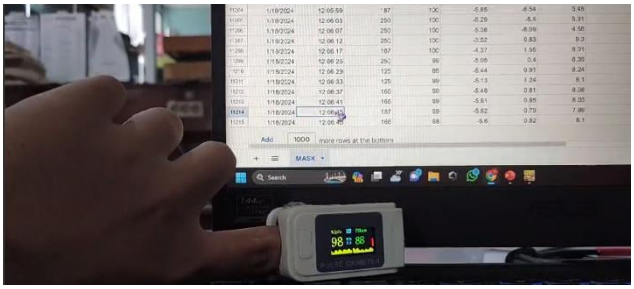


Figure 7. Data validation using Oxymeter.

3.2. Motions Test

The same unit as above, motion test uses MPU6050 as the sensor and transfer the data to microcontroller and then it forwarded to database in the form of x, y, and z axis value and determine the motion of subject's hand which wear the device. Many motions detected in sleeps considered as a sleep disorder where it can disrupt sleep quality[20]. The measurement shown in Fig. 10 shows almost no movement from 3 axises of subject 1 and 2, except for movement at early sleep and subject 3 shows more movement than 2 other subjects. 3 subjects move rarely according to data reading and can be inferred as no disruption to sleep quality [8].

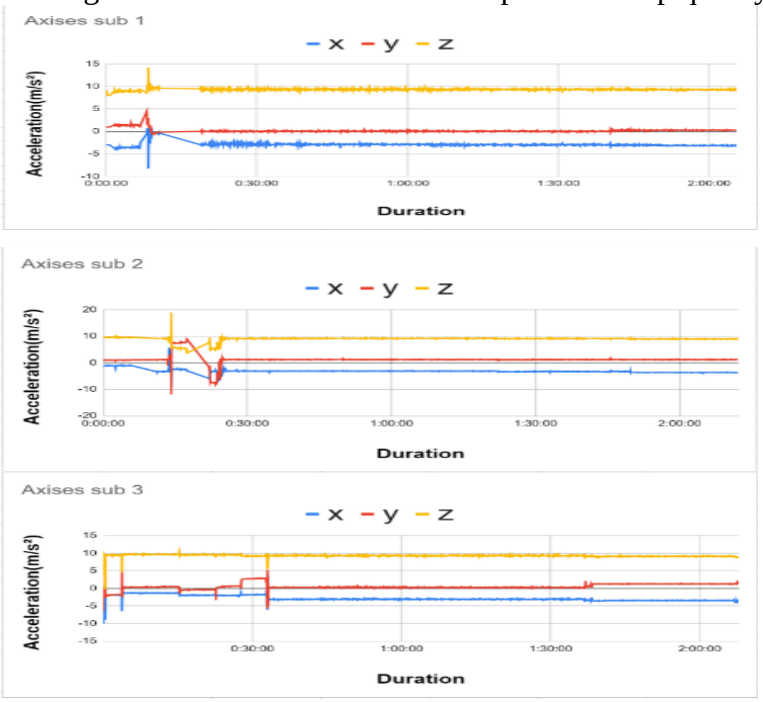


Figure 8. X, Y, and Z axises Measurement from MPU6050.

3.3. Breathing and Sound Test

This test consists of sound data from microphone of MAX9814 and temperature data from thermometer sensor of D18B20. The sound data just came in raw input, the louder the sound is, the sensor will read it as higher value. In Fig. 11, the sounds detected can shows loud

noises especially over 200 and as shown in Fig. 11, the loud sounds can disrupt the sleep quality [8].

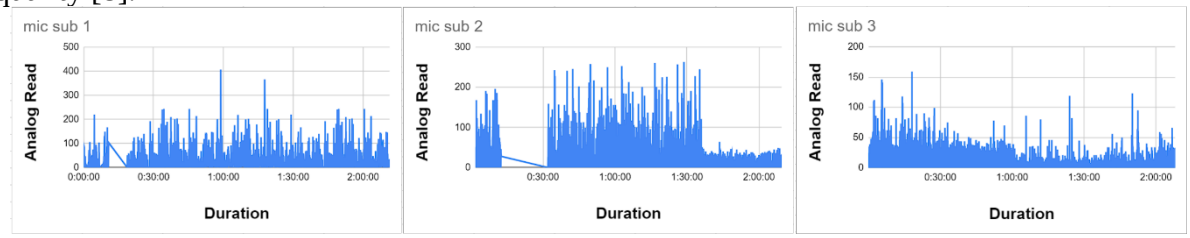


Figure 9. Measurement from MAX9814.

Breathing test uses thermometer to determine the pace of breathing of subject from the changes in value of sensors in reading the heat from breath. Shown in Fig. 11, the changes not really big and will stay around the same value if the pace of breathing is stable. The big changes in value show a change in pace of breathing as well but can't be determined if it change to rapidly taking a breath or stopping for a while.

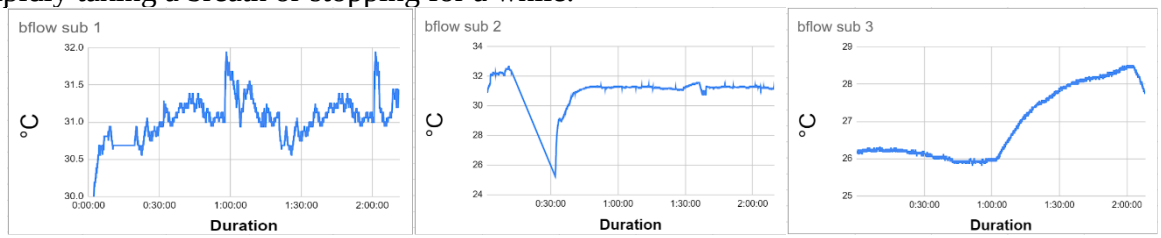


Figure 10. Measurement from DS18B20.

4. Conclusion

Sleeping disorders have many kinds and all types of symptoms. Early detection from these symptoms can prevent other illnesses to struck after countless night of bad quality sleeps. A device to address this issue to detects early symptoms before taking the matter to a professional or a doctor for more advance examination has been developed. The device consisted of 2 microcontroller and 4 sensors of MAX30102 oxygen levels sensors, MPU6050 accelerometer, MAX9814 microphone module, and DS18B20 thermometer. The experiment was conducted 3 times with same environment of sleep and 3 individuals. The results still shown signs of unstable reading which can be improved in the future by improvement in sensors qualities, more experiment conducted, and coordination with professional to calculate and compare more data to get higher accuracy and stable reading. Improvement of this caliber cannot be done as limitation of time and subject, expert to consult, and the objective itself to create affordable home sleep monitor became the main of many reasons.

References

[1] Y. Liu, A. G. Wheaton, D. P. Chapman, T. J. Cunningham, H. Lu, and J. B. Croft, "Prevalence of Healthy Sleep Duration among Adults," *Lead Vis. Inf. Spec.*, vol. 65, no. 6, pp. 137–141, 2016, [Online]. Available: http://www.cdc.gov/

[2] L. Exelmans and J. Van den Bulck, "Sleep Research: A Primer for Media Scholars," *Health Commun.*, vol. 34, no. 5, pp. 519–528, 2019.

[3] Q. Pan, D. Brulin, and E. Campo, "Current Status and Future Challenges of Sleep Monitoring Systems: Systematic Review," *JMIR Biomed. Eng.*, vol. 5, no. 1, p. e20921, 2020, doi: 10.2196/20921.

[4] S. Zarrabian *et al.*, "Structured Clinical Interview for DSM-5—Research Version (SCID-5 for DSM-5, Research Version; SCID-5-RV)," *Drug Alcohol Depend.*, vol. 33, no. 1, pp. 90–106, 2018, [Online]. Available: http://dx.doi.org/10.1038/npp.2017.158%0Ahttps://doi.org/10.1016/j.brainres.2019.02.

026%0Ahttps://academic.oup.com/sleep/article-lookup/doi/10.5665/sleep.3142%0Ahttp://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=1

[5] A. S. Bahammam, E. Alanbay, N. Alrajhi, and A. H. Olaish, "The success rate of split-night polysomnography and its impact on continuous positive airway pressure compliance," *Ann. Thorac. Med.*, vol. 10, no. 4, pp. 274–278, 2015.

[6] I. M. Rosen *et al.*, "Clinical use of a home sleep apnea test: An updated American academy of sleep medicine position statement," *J. Clin. Sleep Med.*, vol. 14, no. 12, pp. 2075–2077, 2018, doi: 10.5664/jcsm.7540.

[7] L. In and S. Quiz, "Sleep Apnea What it is , its risk factors , its health impacts , and how it can be treated What Is Sleep Apnea ?," pp. 1–13, 2023.

[8] National Sleep Foundation, "Managing Excessive Daytime Sleepiness: Its causes and consequences, and how it's diagnosed and treated," pp. 1–8, 2023.

[9] F. M. Ter Heege *et al.*, "The clinical relevance of early identification and treatment of sleep disorders in mental health care: Protocol of a randomized control trial," *BMC Psychiatry*, vol. 20, no. 1, pp. 1–10, 2020, doi: 10.1186/s12888-020-02737-3.

[10] J. G. Claudino *et al.*, "Which parameters to use for sleep quality monitoring in team sport athletes? A systematic review and meta-analysis," *BMJ Open Sport Exerc. Med.*, vol. 5, no. 1, pp. 1–13, 2019, doi: 10.1136/bmjsem-2018-000475.

[11] Einstronic, "Wemos D1 Mini," no. July, pp. 1–8, 2017.

[12] R. C. Cholakian, "Narrative structure in rabelais and the question of the authenticity of the cinquieme livre," *Fr. Stud.*, vol. 33, no. 1, pp. 1–12, 1979, doi: 10.1093/fs/XXXIII.1.1.

[13] A. K. Reimers, V. Heidenreich, H. J. Bittermann, G. Knapp, and C. D. Reimers, "Accelerometer-measured physical activity and its impact on sleep quality in patients suffering from restless legs syndrome," *BMC Neurol.*, vol. 21, no. 1, pp. 1–10, 2021, doi: 10.1186/s12883-021-02115-w.

[14] K. Saleem, I. S. Bajwa, N. Sarwar, W. Anwar, and A. Ashraf, "IoT Healthcare: Design of Smart and Cost-Effective Sleep Quality Monitoring System," *J. Sensors*, vol. 2020, 2020, doi: 10.1155/2020/8882378.

[15] Dallas, "DS18B20 Programmable Resolution 1-Wire Digital Thermometer," *Datasheet DS18B20*, pp. 1–27, 2018.

[16] T. DeBell, L. Goertzen, L. Larson, W. Selbie, J. Selker, and C. Udell, "OPEnS hub: Real-time data logging, connecting field sensors to google sheets," *Front. Earth Sci.*, vol. 7, no. May, pp. 1–6, 2019, doi: 10.3389/feart.2019.00137.

[17] G. Basaranoglu, M. Bakan, T. Umutoglu, S. U. Zengin, K. Idin, and Z. Salihoglu, "Comparison of SpO2 values from different fingers of the hands," *Springerplus*, vol. 4, no. 1, pp. 2–4, 2015, doi: 10.1186/s40064-015-1360-5.

[18] World Health Organization, "Pulse Oximetry Training Manual," *World Heal. Organ.*, pp. 7–8, 2011.

[19] J. Fu and N. Wang, "A Practical Attribute-Based Document Collection Hierarchical Encryption Scheme in Cloud Computing," *IEEE Access*, vol. 7, no. 5, pp. 36218–36232, 2019, doi: 10.1109/ACCESS.2019.2905346.

[20] G. U. I. De, "Healthy Sleep.," *Lancet*, vol. 216, no. 5578, pp. 198–199, 2005, doi: 10.1016/S0140-6736(01)09093-6.

Journal ASSET

Advance Sustainable Science, Engineering and Technology

Publisher: Science and Technology Research Centre
Universitas PGRI Semarang INDONESIA

Address : Jl. Lontar No. 1 Semarang
Central Java 50232 INDONESIA

E-Mail : asset@upgris.ac.id

ISSN: 2715-4211