Fast and Accurate Indonesian QnA Chatbot Using Bag-of-Words and Deep-Learning for Car Repair Shop Customer Service

Muchamad Taufiq Anwar[1*], Azzahra Nurwanda[1], Fajar Rahmat[1], Muhammad Aufal[1], Hindriyanto Dwi Purnomo[2], Aji Supriyanto[3]

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

[2]Faculty of Information Technology, Universitas Kristen Satya Wacana, Jl. Dr. O. Notohamidjodjo, Salatiga 50715, Central Java, Indonesia

[3]Faculty of Information Technology, Universitas Stikubank, Jl. Tri Lomba Juang No 1 Semarang 50241, Central Java, Indonesia

*taufiq@stmi.ac.id

Abstract. A chatbot is a software that simulates human conversation through a text chat. Chatbot is a complex task and recent approaches to Indonesian chatbot have low accuracy and are slow because it needs high resources. Chatbots are expected to be fast and accurate especially in business settings so that they can increase customer satisfaction. However, the currently available approach for Indonesian chatbots only has low to medium accuracy and high response time. This research aims to build a fast and accurate chatbot by using Bag-of-Words and Deep-Learning approach applied to a car repair shop customer service. Sixteen different intents with a set of their possible queries were used as the training dataset. The approach for this chatbot is by using a text classification task where intents will be the target classes and the queries are the text to classify. The chatbot response then is based on the recognized intent. The deep learning model for the text classification was built by using Keras and the chatbot application was built using the Flask framework in Python. Results showed that the model is capable of giving 100% accuracy in predicting users' intents so that the chatbot can give the appropriate responses and the response time is near zero milliseconds. This result implies that developers who aim to build fast and accurate chatbot software can use the combination of bag-of-words and deep-learning approaches. Several suggestions are presented to increase the probability of the chatbot's success when released to the general public.

Keywords: Indonesian chatbot, car repair shop customer service, fast, accurate, deep learning, bag of words

1. Introduction

A chatbot is a software that simulates human conversation through a text chat. It is one of the tasks in Natural Language Processing (NLP) in the field of Artificial Intelligence. Chatbots are expected to be fast and accurate [1] and chatbot performance is important to increase customer experience and therefore improve customer satisfaction [2]. However, the currently available approach for Indonesian chatbots only has low to medium accuracy and high response time. The accuracy of the chatbot is a measure of how likely the chatbot is to give the correct (matching) answer to the question asked by the user [3]. Whereas response time is defined as the time required by the chatbot to give a response based on the user's query [4]. Research [5] has an accuracy of 26-78% using pre-built platforms namely Artificial Intelligence Markup Language (AIML) and Pandorabots. Another research has an accuracy of up to 93.75% [6] and another research using RNN for intent classification has an accuracy of 81% [7]. Another research using the GloVe model and CNN produces the best accuracy of 95.84% [8]. Only one research reports the chatbot's response time and it has a low response time with an average response time of 3.4 seconds [4]. Chatbot in Bahasa Indonesia had been used for FAQ in a university library during a pandemic situation [9], virtual assistant for banking service [10], [11], hotel [12], [13], software requirements elicitation [14], legal information service [15], digital historian [16], maternal and child health [17], and e-commerce [5].

The approach to building a chatbot varies from traditional methods such as pattern matching [11] and cosine similarity [18], to the more modern, neural-network-based approaches such as sequence-to-sequence model [19] and Transformers[20]. However, it has been proved that the development of a pattern-matching or rule-based approach is very difficult [21]. The later approaches will also allow the users to use free text as their queries, rather than a fixed pattern [22]. Some authors may use the pre-build platform such as RASA [23], Dialogflow [9], and AIML [5], [12]–[14]. This research aims to build a deep-learning-based chatbot from scratch using TensorFlow/Keras.

A deep neural network is a type of neural network that has more than one hidden layer. The higher complexity of the deep neural network will allow the model to do a more complex task including language modeling and Question-and-Answering task such as chatbot. For a computer to be able to process text, the text needs to be pre-processed and undergo feature extraction steps before being fed into the model. The text pre-processing tasks include tokenization, lemmatization, stemming, stop-word removal, etc. There are some text feature extraction techniques such as (Term Frequency–Inverse Document Frequency) TF-IDF [24], Term Document Matrix (TDM) [25], Bag of Words (BoW) [26], Word Embeddings [27], etc. Research showed that when using BoW, the text conversion of uppercase letters into lowercase letters gave a significant improvement in text classification accuracy [26]. Whereas stop word removal may improve the accuracy but the result is dependent on the dataset used [26]. The result from text feature extraction is a matrix that later is used for text classification. The text classification part may use some "traditional" Data Mining techniques such as Random Forest and Gradient Boosting [24] or a more modern, neural-network-based approach, such as Transformers. The text classification task may use Transformers' zero-shot learning approach but it is slow / resource-intensive [28].

Customer Service (CS) is one of the important parts of every business, including a car repair shop. Traditionally, CS tasks are held by human CS. However, when so many customers only ask about basic information, this can burden the utilization of human CS where the precious resource of human CS are wasted on answering very basic questions from the customers. The use of a chatbot can help CS [1] to alleviate this burden and reduce human resource costs for CS [29]. Furthermore, the use of chatbots can also avoid human error [29] and provide the latest information about the business [1]. Regardless of being a non-human CS, chatbots are proven to increase business agility [30], form positive customer relationships [31], and increase customer satisfaction [2], [32], [33]. This research will apply the chatbot to help a car repair shop CS. If the customers need further questions, the chatbot will direct the customer to a human CS.

2. Methods

The research methods are shown in Figure 1 which shows that the methods are separated into two phases: the training phase and the implementation phase. In the training phase, a set of preset queries and their corresponding intent is prepared as the training dataset in the JSON file format. The queries consist of questions such as "What is the opening hour?" and "How long will the repair take time?". There are a total of 16 intents with each intent having several possible queries. There is one special intent to handle the case when the user inputs an invalid or blank query or queries that the chatbot can not understand. This special intent will be responded to by the chatbot asking the user to type a valid / better query. The approach for this chatbot is by using a text classification task where intents will be the target classes and the queries are the text to classify. The chatbot response then is based on the recognized intent. An example of the training data is shown in Figure 2.

The preset queries will undergo text preprocessing and feature extraction steps which includes tokenization, lemmatization / stemming, and bag-of-words creation. Tokenization will list all words in the query. Lemmatization and stemming will change the words into their root form. Bag of Words will create a corpus that lists all unique words and count the word frequency for each word. These steps will result in a matrix that is ready to be used for model training. The model consists of one input layer, two hidden layers, and one output layer. The details of the model architecture are shown in Table 2. The model is trained using SGD optimizer with Learning Rate = 0.1. The trained model will be saved as an H5 file to be further used in the chatbot application development.

In the deployment phase, the user will input queries in the app which will then be replied to by the chatbot. The query will undergo the same text preprocessing and feature extraction as in the training phase before being fed into the model. The model will then output the predicted intent and the app will give the appropriate response based on the predicted intent. The confidence threshold for the intent prediction is set to 0.25 and the model will give a response based on the class with the highest confidence score. If there is more than one preset response, the app will give a random preset response for that intent. In case of a low confidence prediction score, the app will respond with a statement that the chatbot does not understand the user query. And for special or further queries, the chatbot will direct the customer to human customer service. Both the training and deployment phases are done in the Python environment by using several libraries such as NLTK [34] for tokenization and lemmatization, Sastrawi for Indonesian stopword removal and stemming, Tensorflow [35] and Keras [36] for the deep learning model, Pickle for saving the model, and Flask [37] as the framework to develop the web-based chatbot app.
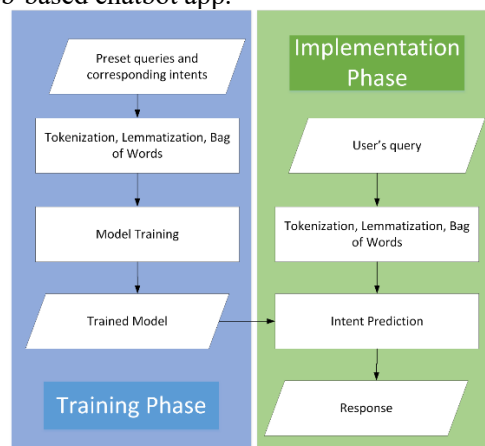

Figure 1. Research Methods

```
{"tag": "1",
 "patterns": ["Jam Operasional","Jam operasional
 bengkel?","Apakah bengkel buka?","Buka jam berapa?",
 "Tutup jam berapa?","Jam Buka"],
 "responses": ["Bengkelin melayani servis mobil
 setiap hari, kecuali hari libur. Kami buka mulai
 pukul 08.00 - 20.00 WIB. Silahkan berkunjung
 \uD83D\uDE0A",
                        "Halo kak, jam operasional kami
                        dimulai dari pukul 08.00 - 20.00 WIB
                        ya",
                        "Silahkan datang ke bengkel kami pada
                        pukul 08.00 - 20.00 WIB"],
 "context": [""]
},
```

Figure 2. A sample of intent, its preset queries, and its preset response in the JSON format

Table 1. Model architecture

Layer	Details
Input layer	Input shape equals the number of BoW elements
Hidden layer 1	128 neurons, dropout rate = 0.5, ReLu AF
Hidden layer 2	64 neurons, dropout rate = 0.5, ReLu AF
Output layer	Neuron units = number of intents, Softmax AF

3. Results and Discussion

By using 200 epochs, the training phase resulted in a model with 100% accuracy as shown in Figure 3. The training time was two minutes. The trained model is saved and then deployed in a web-based Python app. The result of the app is shown in Figure 4 and Figure 5. User will input their queries in the text field and hit the "Send" button and the chatbot will respond to it. For example, when the user asked about facilities, the chatbot will respond with information regarding the available facilities in the car repair shop. The timestamp of each message is displayed in each of the bubble chats. Figure 5 shows that when users' queries contain abbreviated words and/or typos (writing errors), the chatbot can still understand them and give the appropriate response (as long as another/other important keywords are present). When the words in the query are all abbreviated and/or typos, the chatbot will return a response saying that it can not understand the query, as shown in Figure 6. Figure 7 shows the Python console showing the difference between the user's query time and chatbot response time. The difference is near zero milliseconds. This shows that our results is better in both accuracy and speed than previous research mentioned in the Introduction section. The use of this Chatbot will certainly help the car repair shop CS to serve their customers who want to ask questions about their services. CS will only focus on serving customers who need further information about its services.

Figure 3. The trained model has 100% accuracy
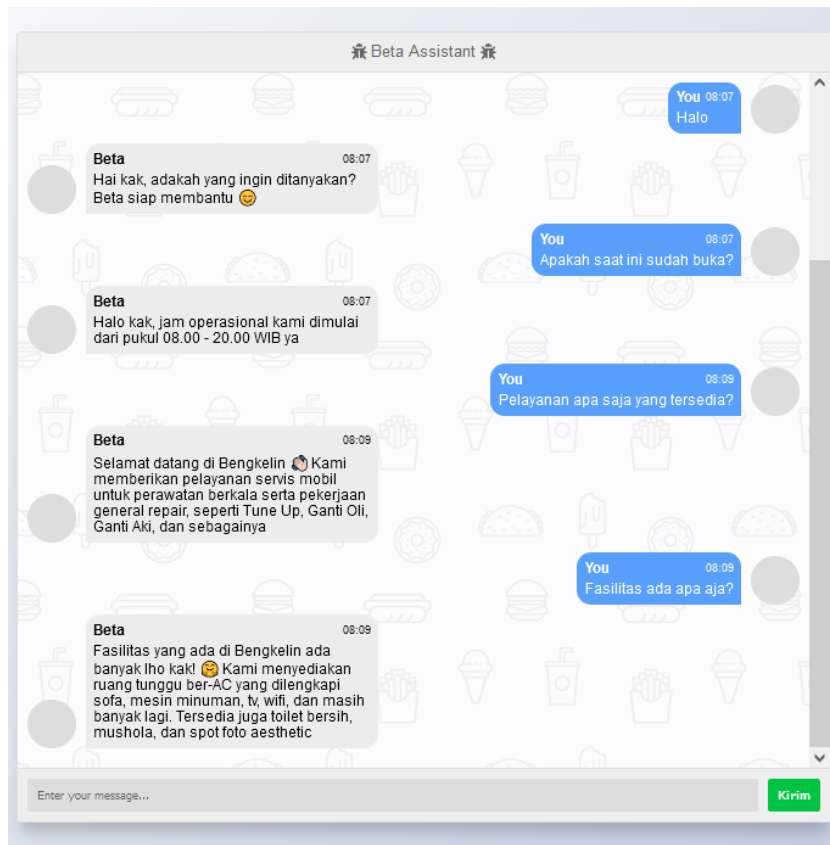

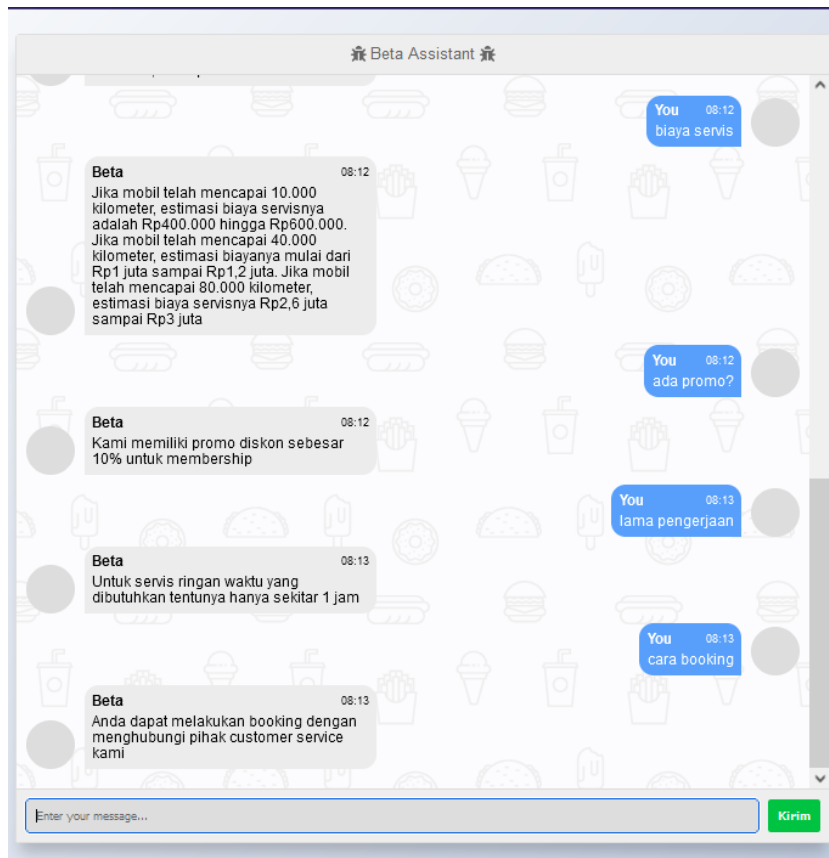**Figure 4.** The result of the chatbot app

Figure 5. A screenshot of the chatbot app showing its capability to respond to incomplete/short sentences
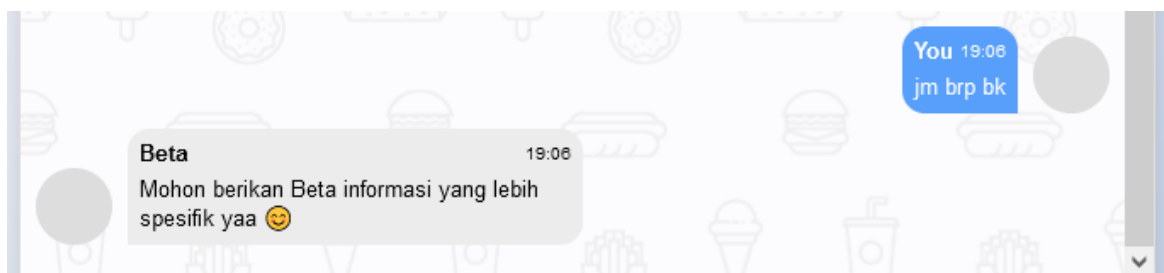

Figure 6. A screenshot of the chatbot app showing a respond to a sentence containing all abbreviated words and typos. The chatbot will respond by asking the user to input a more detailed / better-typed query

Despite its great accuracy and speed, when released to the public, research showed that a successful chatbot needs to be useful and easy to use [38], [39]. Research also showed that user satisfaction in using a chatbot is influenced by System Quality, Information Quality, and Service Quality [40]. Whereas the intention to use a chatbot is influenced by Information Quality and Personal factors (such as age and occupation) [40]. Further research may address and implement those insights to guarantee a successful chatbot implementation that is well-received by the public. Further research may also use a more sophisticated out-of-scope intent detection [20] rather than by providing one extra intent class as used in this research. Latent Dirichlet Allocation (LDA), a text clustering technique [41], might be explored to be used as an approach to / to assist the text classification task [42], [43]. Future research may also add Typographical Error correction to make the chatbot more robust in handling Typographical Errors using a method such as Schema Matching Technique [44]. Future research may add more datasets for both preset queries and their corresponding intents to further expand the functionality of the chatbot. The chatbot can be further improved by integrating the chatbot into the car repair shop information system so that it can handle service orders, provide real-time information about the current service queue, provide real-time information on the estimated time to complete service, etc. But such system may require a referee policy so that it will not be abused by irresponsible users.

Figure 7. The Python console showing the difference between the user's query time and chatbot response time. The difference is near zero milliseconds

4. Conclusion
This research aims to build a fast and accurate Indonesian QnA chatbot using bag-of-words and deep-learning approach applied to a car repair shop. The training phase resulted in a 100% accurate model. Results showed that the chatbot app can provide a fast (near-zero milliseconds response time) and highly accurate response based on users' queries. This chatbot can help to reduce the cost use of hiring human CS and reduce human errors. Despite from its already great performance, there are several suggestions to increase its chance of success when released to the general public.

References

[1]   A. C. Sari, N. Virnilia, J. T. Susanto, K. A. Phiedono, and T. K. Hartono, "Chatbot developments in the business world," *Adv. Sci. Technol. Eng. Syst. J.*, vol. 5, no. 6, pp. 627–635, 2020.

[2]   I. Rizomyliotis, M. N. Kastanakis, A. Giovanis, K. Konstantoulaki, and I. Kostopoulos, "'How mAy I help you today?' The use of AI chatbots in small family businesses and the moderating role of customer affective commitment," *J. Bus. Res.*, vol. 153, pp. 329–340, 2022.

[3]   G. Aalipour, P. Kumar, S. Aditham, T. Nguyen, and A. Sood, "Applications of sequence to sequence models for technical support automation," in *2018 IEEE International Conference on Big Data (Big Data)*, 2018, pp. 4861–4869.

[4]   A. Nursetyo, E. R. Subhiyakto, and others, "Smart chatbot system for E-commerce assitance

based on AIML," in *2018 International Seminar on Research of Information Technology and Intelligent Systems (ISRITI)*, 2018, pp. 641–645.

[5]   S. A. Prasetya, A. Erwin, and M. Galinium, "Implementing Indonesian language chatbot for ecommerce site using artificial intelligence markup language (AIML)," in *Prosiding Seminar Nasional Pakar*, 2018, pp. 313–322.

[6]   M. A. Hakim, S. Nurhayati, and K. Indonesia, "Development of Chatbot Application 'Midwify 'Based on Android As a Supporting Media To Learn Medical Science in Stikes Bhakti Kencana Bandung," *Komputika J. Sist. Komput.*, vol. 8, no. 1, 2019.

[7]   A. Prasetyo and H. A. Santoso, "Intents Categorization for Chatbot Development Using Recurrent Neural Network (RNN) Learning," in *2021 7th International Conference on Advanced Computing and Communication Systems (ICACCS)*, 2021, vol. 1, pp. 51–55.

[8]   C. O. Bilah, T. B. Adji, and N. A. Setiawan, "Intent Detection on Indonesian Text Using Convolutional Neural Network," in *2022 IEEE International Conference on Cybernetics and Computational Intelligence (CyberneticsCom)*, 2022, pp. 174–178.

[9]   S. P. Barus and E. Surijati, "Chatbot with Dialogflow for FAQ Services in Matana University Library," *Int. J. Informatics Comput.*, vol. 3, no. 2, pp. 62–70, 2022.

[10]   B. P. Wicaksono and A. Zahra, "Design of the use of chatbot as a virtual assistant in banking services in Indonesia," *IAES Int. J. Artif. Intell.*, vol. 11, no. 1, p. 23, 2022.

[11]   A. Elcholiqi and A. Musdholifah, "Chatbot in Bahasa Indonesia using NLP to provide banking information," *IJCCS (Indonesian J. Comput. Cybern. Syst.*, vol. 14, no. 1, pp. 91–102, 2020.

[12]   D. Gunawan, F. P. Putri, and H. Meidia, "Bershca: bringing chatbot into hotel industry in Indonesia," *TELKOMNIKA (Telecommunication Comput. Electron. Control.*, vol. 18, no. 2, pp. 839–845, 2020.

[13]   F. P. Putri, H. Meidia, and D. Gunawan, "Designing intelligent personalized chatbot for hotel services," in *Proceedings of the 2019 2nd International Conference on Algorithms, Computing and Artificial Intelligence*, 2019, pp. 468–472.

[14]   F. Dwitama and A. Rusli, "User stories collection via interactive chatbot to support requirements gathering," *TELKOMNIKA (Telecommunication Comput. Electron. Control.*, vol. 18, no. 2, pp. 890–898, 2020.

[15]   V. A. H. Firdaus, P. Y. Saputra, and D. Suprianto, "Intelligence chatbot for Indonesian law on electronic information and transaction," in *IOP Conference Series: Materials Science and Engineering*, 2020, vol. 830, no. 2, p. 22089.

[16]   K. T. Wirawan, I. M. Sukarsa, and I. P. A. Bayupati, "Balinese historian chatbot using full-text search and artificial intelligence markup language method," *Int. J. Intell. Syst. Appl.*, vol. 11, no. 8, p. 21, 2019.

[17]   I. W. Puspitasari, F. R. Rinawan, W. G. Purnama, H. Susiarno, and A. I. Susanti, "Development of a Chatbot for Pregnant Women on a Posyandu Application in Indonesia: From Qualitative Approach to Decision Tree Method," in *Informatics*, 2022, vol. 9, no. 4, p. 88.

[18]   L. Hidayatin and F. Rahutomo, "Query expansion evaluation for chatbot application," in *2018 International Conference on Applied Information Technology and Innovation (ICAITI)*, 2018, pp. 92–95.

[19]   Y. W. Chandra and S. Suyanto, "Indonesian chatbot of university admission using a question answering system based on sequence-to-sequence model," *Procedia Comput. Sci.*, vol. 157, pp. 367–374, 2019.

[20]   L. Manik *et al.*, "Out-of-scope intent detection on a knowledge-based chatbot," *Int. J. Intell. Eng. Syst.*, vol. 14, no. 5, pp. 446–457, 2021.

[21]   L. Shang, Z. Lu, and H. Li, "Neural responding machine for short-text conversation," *arXiv Prepr. arXiv1503.02364*, 2015.

[22]   C. Pricilla, D. P. Lestari, and D. Dharma, "Designing interaction for chatbot-based conversational commerce with user-centered design," in *2018 5th International Conference on*

Advanced Informatics: Concept Theory and Applications (ICAICTA), 2018, pp. 244–249.

[23] Y. Windiatmoko, R. Rahmadi, and A. F. Hidayatullah, "Developing facebook chatbot based on deep learning using rasa framework for university enquiries," in *IOP conference series: materials science and engineering*, 2021, vol. 1077, no. 1, p. 12060.

[24] M. T. Anwar, "Automatic Complaints Categorization Using Random Forest and Gradient Boosting," *Adv. Sustain. Sci. Eng. Technol.*, vol. 3, no. 1, p. 210106, 2021.

[25] E. Zuliarso, M. T. Anwar, K. Hadiono, and I. Chasanah, "Detecting Hoaxes in Indonesian News Using TF/TDM and K Nearest Neighbor," in *IOP Conference Series: Materials Science and Engineering*, 2020, vol. 835, no. 1, p. 12036.

[26] Y. HaCohen-Kerner, D. Miller, and Y. Yigal, "The influence of preprocessing on text classification using a bag-of-words representation," *PLoS One*, vol. 15, no. 5, p. e0232525, 2020.

[27] P. L. Rodriguez and A. Spirling, "Word embeddings: What works, what doesn't, and how to tell the difference for applied research," *J. Polit.*, vol. 84, no. 1, pp. 101–115, 2022.

[28] 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.

[29] M. Ridha and K. Haura Maharani, "Implementation of Artificial Intelligence Chatbot in Optimizing Customer Service in Financial Technology Company PT. FinAccel Finance Indonesia," *Multidiscip. Digit. Publ. Inst. Proc.*, vol. 83, no. 1, p. 21, 2022.

[30] X. Wang, X. Lin, and B. Shao, "How does artificial intelligence create business agility? Evidence from chatbots," *Int. J. Inf. Manage.*, vol. 66, p. 102535, 2022.

[31] M. Chung, E. Ko, H. Joung, and S. J. Kim, "Chatbot e-service and customer satisfaction regarding luxury brands," *J. Bus. Res.*, vol. 117, pp. 587–595, 2020.

[32] Y. Ruan and J. Mezei, "When do AI chatbots lead to higher customer satisfaction than human frontline employees in online shopping assistance? Considering product attribute type," *J. Retail. Consum. Serv.*, vol. 68, p. 103059, 2022.

[33] A. Wiliam, S. Sasmoko, H. Prabowo, and others, "Analysis of e-service chatbot and satisfaction of banking customers in Indonesia," *Asia Proc. Soc. Sci.*, vol. 4, no. 3, pp. 72–75, 2019.

[34] S. Bird, E. Klein, and E. Loper, *Natural language processing with Python: analyzing text with the natural language toolkit*. " O'Reilly Media, Inc.," 2009.

[35] Martín~Abadi *et al.*, "{TensorFlow}: Large-Scale Machine Learning on Heterogeneous Systems." 2015. [Online]. Available: https://www.tensorflow.org/

[36] F. Chollet and others, "Keras." 2015.

[37] M. Grinberg, *Flask web development: developing web applications with python*. " O'Reilly Media, Inc.," 2018.

[38] R. Richad, V. Vivensius, S. Sfenrianto, and E. R. Kaburuan, "Analysis of factors influencing millennial's technology acceptance of chatbot in the banking industry in Indonesia," *Int. J. Civ. Eng. Technol.*, vol. 10, no. 4, pp. 1270–1281, 2019.

[39] L. Sanny, A. Susastra, C. Roberts, and R. Yusramdaleni, "The analysis of customer satisfaction factors which influence chatbot acceptance in Indonesia," *Manag. Sci. Lett.*, vol. 10, no. 6, pp. 1225–1232, 2020.

[40] J. A. Mulyono and S. Sfenrianto, "Evaluation of customer satisfaction on Indonesian banking chatbot services during the COVID-19 pandemic," *CommIT (Communication Inf. Technol. J.*, vol. 16, no. 1, pp. 69–85, 2022.

[41] 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.

[42] J. Jedrzejowicz and M. Zakrzewska, "Text classification using LDA-W2V hybrid algorithm,"

in *Intelligent Decision Technologies 2019: Proceedings of the 11th KES International Conference on Intelligent Decision Technologies (KES-IDT 2019), Volume 1*, 2020, pp. 227–237.

[43] M. Xian-yan, C. Rong-yi, Z. Ya-hui, and Z. Zhenguo, "Multilingual short text classification based on LDA and BiLSTM-CNN neural network," in *Web Information Systems and Applications: 16th International Conference, WISA 2019, Qingdao, China, September 20-22, 2019, Proceedings 16*, 2019, pp. 319–323.

[44] M. L. Tedjopranoto, A. Wijaya, L. H. Santoso, and D. Suhartono, "Correcting typographical error and understanding user intention in Chatbot by combining N-gram and machine learning using schema matching technique," *Int. J. Mach. Learn. Comput.*, vol. 9, no. 4, pp. 471–476, 2019.