

# PENERAPAN ALGORITMA K-MEANS CLUSTERING PADA DATA KELUHAN PELANGGAN PT. PLN PERSERO KOTA PAGAR ALAM

Sasmita<sup>1)</sup>, Siti muntari<sup>2)</sup>

<sup>1)</sup>Prodi Teknik Informatika, Institut Teknologi Pagar Alam, Jl Masik Siagim no. 75 Simpang Mbacang Dempo Tengah Pagar Alam

<sup>2)</sup> Prodi Teknik Informatika, Institut Teknologi Pagar Alam, Jl Masik Siagim no. 75 Simpang Mbacang Dempo Tengah Pagar Alam

Email: [sasmitha661@gmail.com](mailto:sasmitha661@gmail.com), [muntariaza@gmail.com](mailto:muntariaza@gmail.com)

## Abstrak

Di zaman modern ini semua aktivitas serta kebutuhan warga sebagian besar sangat di pengaruhi oleh listrik. Listrik di butuhkan sebab semua peralatan rumah tangga menggunakan tenaga listrik buat kebutuhan perusahaan atau kebutuhan tempat tinggal. Untuk meningkatkan kualitas pelayanan PT. PLN Persero Kota Pagar Alam agar menekan banyaknya keluhan pelanggan di Kota Pagar Alam dibutuhkan sebuah proses pengelompokan data (data *Clustering*) yang sangat penting di karenakan meningkatnya data yang cukup signifikan. Proses pengelompokan data menggunakan *K-Means Clustering* dikarenakan Algoritma ini sesuai untuk mengelompokkan data tersebut. Hasil penelitian ini Berupa perancangan penerapan algoritma *K-Means Clustering* pada data keluhan pelanggan PT. PLN Persero Kota Pagar Alam dimana *cluster* terbagi menjadi 3 yaitu C0 sebagai tingkat keluhan tertinggi, C1 tingkat keluhan sedang dan C2 tingkat keluhan rendah. Proses import data menggunakan *rapidminer*. Dari pola yang didapatkan di *rapidminer* yang digunakan pada sistem dengan metode *clustering k-means* maka diperoleh bahwa *cluster\_0* memiliki tingkat keluhan tinggi dengan jumlah 108, *cluster\_1* memiliki tingkat keluhan sedang dengan jumlah 75 dan *cluster\_2* memiliki tingkat keluhan rendah dengan jumlah 45.

**Kata Kunci :** PT. PLN, Persero, K-Means, Clustering, Rapid Application Development

## Abstract

*In this modern era, all activities and needs of residents are largely influenced by electricity. Electricity is needed because all household appliances use electric power for company needs or residential needs. To improve the service quality of PT. PLN Persero, Pagar Alam City, in order to reduce the number of customer complaints in Pagar Alam City, a data clustering process is needed which is very important because the increase in data is quite significant. The process of grouping data uses K-Means Clustering because this algorithm is suitable for grouping the data. The results of this study are in the form of designing the application of the K-Means Clustering algorithm to customer complaint data at PT. PLN Persero City of Pagar Alam where the cluster is divided into 3, namely C0 as the highest complaint level, C1 for moderate complaint levels and C2 for low complaint levels. The data import process uses rapidminer. From the pattern obtained in rapidminer which is used on the system with the k-means clustering method, it is obtained that cluster\_0 has a high complaint rate with a total of 108, cluster\_1 has a moderate complaint rate with a total of 75 and cluster\_2 has a low complaint rate with a total 45.*

**Keywords:** PT. PLN, Persero, K-Means, Clustering, Rapid Application Development

## 1. PENDAHULUAN

Di zaman modern ini semua aktivitas serta kebutuhan warga sebagian besar sangat di pengaruhi oleh listrik. Listrik di butuhkan sebab semua peralatan rumah tangga menggunakan tenaga listrik buat kebutuhan perusahaan atau kebutuhan tempat tinggal .(Dahlan et al., 2020)

Listrik adalah kebutuhan yang di anggap krusial. Riset ini menjelaskan mengenai implementasi algoritma K-Means Clustering data keluhan pelanggan. Yang disajikan dalam 3 cluster yaitu cluster keluhan tertinggi (C1), cluster keluhan sedang (C2), serta cluster keluhan rendah (C3). Mengolah data memakai software RapidMiner sebagai penyediaan sehingga memperoleh hasil akhir. Akibat analisis berasal penelitian ini bisa menjadi tuang buat pln supaya membuahkan laporan keluhan pelanggan seperti kategori tertinggi (C1) menjadi hal utama pada menaikkan pelayanan pelanggan dalam menekan banyaknya keluhan pada 2 jenis laporan tertinggi yaitu PDPB dan Prabayar. (perkara et al., 2022)

Contoh Clustering Menggunakan algoritma K-Means mengenai Keluhan Pelanggan PT. PLN. PLN merupakan wadah penyedia jasa pelayanan awam mengawasi sektorketenaga listrikan di Indonesia harus meningkatkan bentuk pelayanan dengan konsumen. Pengelompokan data bertujuan membagi data pada beberapa kelompok (cluster) sehingga mempunyai keadaan yang sama (intra-cluser)

mempunyai derajat kecenderungan yang tinggi, sementara data-data pada cluster yang tidak selaras (intra-cluser) memiliki derajat kesamaan yang rendah. model clustering data gangguan yang pada rancang dengan metode algoritma K-Means, model perangkat lunak mampu menunjukkan gambaran dan memunculkan pola penyebaran data keluhan pelanggan. (Hartanti et al., 2015)

PT. PLN ialah BUMN yang mengurus seluruh aspek kelistrikan di Indonesia. Tujuan PT. PLN yaitu menyelenggarakan usaha penyediaan energi listrik bagi kebutuhan awam. PT. PLN menjadi perusahaan pelayanan listrik terbesar pada Indonesia, PLN mempunyai tantangan yaitu jumlah pengaduan dan keluhan pelanggan yang bertambah seiring meningkatnya jumlah pengguna listrik di Indonesia. (Rohmatullah, 2018)

Sesuai observasi dan wawancara pada PT. PLN (Persero), sistem pengaduan pada PT. PLN telah menggunakan sebuah perangkat lunak yang bernama PLN MOBILE dimana warga atau pelanggan dapat mengadukan keluhan mengenai pencatatan meter (Cater), Pemutusan Sambungan (Tusbung), Pemakaian Listrik tidak sah, Perubahan Daya Pasang Baru (PDPB), rekening, Mutu serta Keandalan, Prabayar dan Instalasi Listrik. akan tetapi pada software tersebut belum adanya proses cluster atau pengelompokan ihwal data keluhan pelanggan. PT. PLN Persero susah buat mendeteksi keluhan masyarakat pada optimalisme pada melakukan penanganan keluhan rakyat.

buat menaikkan kualitas pelayanan PT.PLN Persero Kota Pagar Alam supaya menekan banyaknya keluhan pelanggan di Kota Pagar Alam supaya menekan banyaknya keluhan pelanggan di Kota Pagar Alam diperlukan sebuah proses pengelompokan data (data Clustering) yang begitu diperlukan disebabkan meningkatnya data relatif signifikan. Proses cluster memakai K-Means Clustering dikarenakan Metode ini sesuai untuk mengelompokkan data tadi.

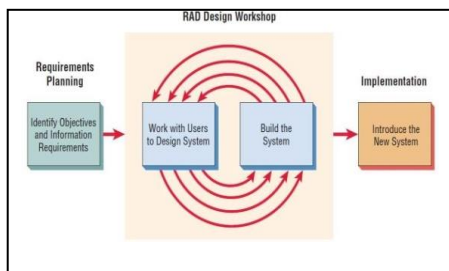
K-means ialah metode data clustering yg berusaha memartisi data yg terdapat kedalam bentuk kelompok sebagai akibatnya data yg mempunyai kesamaan dikelompokkan pada satu cluster yg sama serta data karakteristik yg berbeda dicluster ke dalam cluster yang lainnya.(Metisen & Sari, 2015)

Dari beberapa riset terdahulu bisa disimpulkan bahwa beberapa penelitian di atas bisa menjadi referensi bagi penulis. dengan melihat pertarungan serta berdasarkan uraian di atas, maka dihasilkan suatu judul penelitian **“PENERAPAN ALGORITMA K-MEANS CLUSTERING PADA DATA KELUHAN PELANGGAN PT.PLN PERSERO KOTA PAGAR ALAM”**.

## 2. METODE

### 2.1 Metode Pengembangan Sistem

Metode pengembangan sistem yang digunakan dalam penelitian ini yaitu Rapid Application Development (RAD)



Gambar 2.1 : RAD Design WorkShop

#### 2.1.1 Requirements Planning

Tujuan dari Requirements Planning ini adalah untuk mengidentifikasi kebutuhan, batasan dan alternatif pemecahan masalah dari sistem yang akan di buat.

#### 2.1.2 Workshop Design

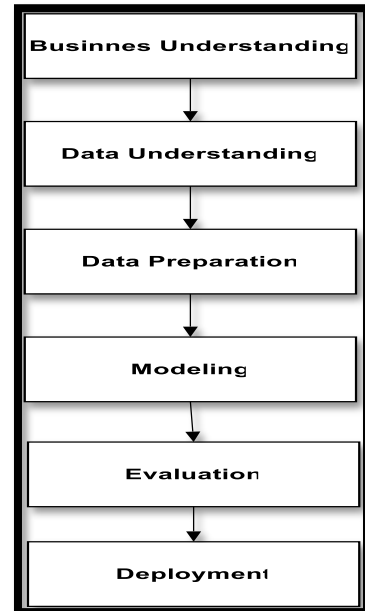
Workshop Design yaitu merancang semua kegiatan arsitektur sistem secara keseluruhan dengan melibatkan abstraksi, identifikasi dan deskripsi sistem perangkat lunak yang mendasar dan hubungan-hubungannya. Tools yang digunakan dalam pemodelan sistem biasanya menggunakan (UML).

### 2.1.3 Implementation

Fase ini dimana tahap pembuatan program penerapan algoritma K-Means Clustering pada data keluhan pelanggan PT.PLN Persero Kota Pagar Alam serta pengujian sistem.

### 2.1.4 Desain Penelitian

Dalam penelitian ini peneliti menggunakan model standarisasi data mining yaitu CRISP-DM (Cross Industry Standart Process for Data Mining), dengan langkah – langkah sebagai berikut :



Gambar 3.2 : Model referensi CRISP-DM

### 2.1.5 Pemahaman Bisnis (Business Understanding)

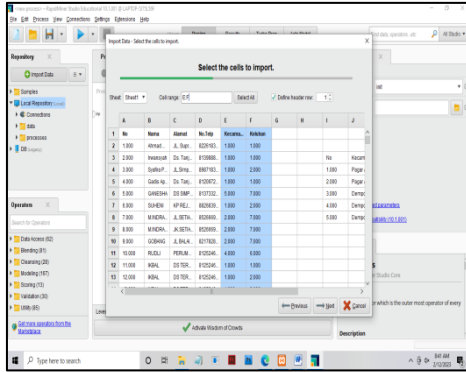
Penelitian ini mengedepankan penentuan tujuan dari algoritma K-means dalam meng cluster dan kebutuhan secara detail dalam lingkup bisnis atau unit penelitian secara keseluruhan, serta juga menerjemahkan tujuan dan batasan dari data yang diambil dari PT.PLN Persero Kota Pagar Alam untuk menjadi formula dari permasalahan data mining mulai dari menyiapkan strategi awal hingga metode yang dibutuhkan untuk mencapai tujuan.

### 3. HASIL DAN PEMBAHASAN

#### 3.1 Hasil

##### 3.1.1. Dataset

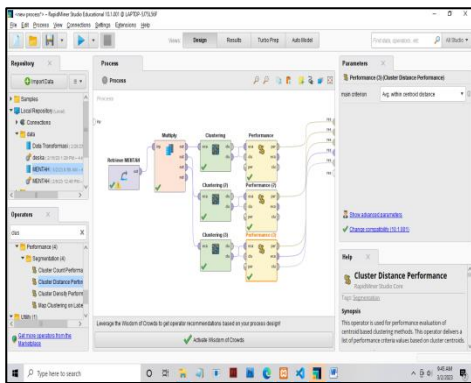
Proses import data menggunakan data keluhan pelanggan PT.Pln Persero Kota Pagar Alam. Dimana dapat dilihat di bawah ini yaitu tampilan sebuah proses Import data pada menggunakan apk Software RapidMinner.



Gambar 3.2 : Import Data

##### 3.1.2. Process K-Means Cluster

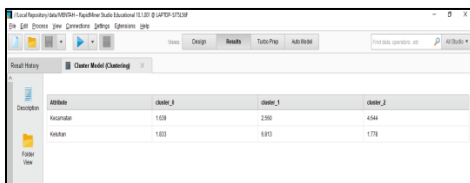
Adapun dimana proses clustering sebagai berikut :



Gambar 3.3 : Process K-Means Cluster

##### 3.1.3. Tentukan Jumlah Cluster

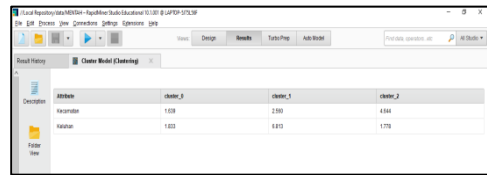
Untuk mengetahui jumlah Cluster dapat dilakukan sebuah percobaan yang kemudian dapat dilihat titik centroidnya.



Gambar 3.5 : Menentukan Jumlah Cluster

##### 3.1.4. Menentukan Titik Centroid Secara Acak

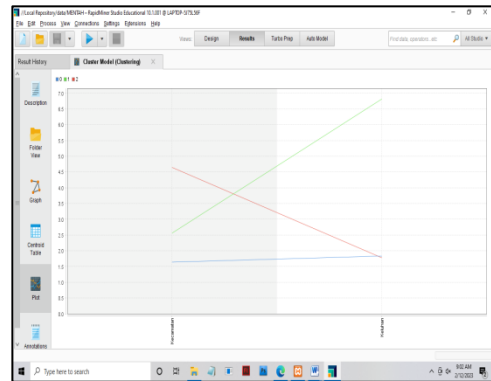
Menentukan titik centroid secara acak berdasarkan percobaan cluster di aplikasi rapidminner. Setelah melakukan percobaan tersebut ditentukanlah titik nilai pusat terkecil untuk mendapatkan hasil yang terbaik.



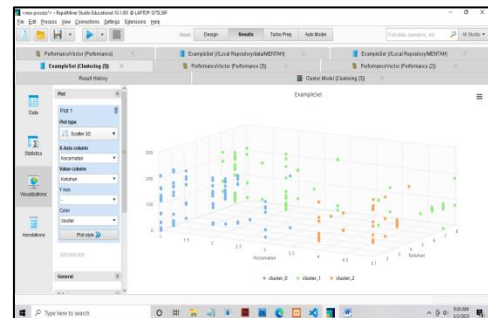
Gambar 3.1 : Titik Centroid Secara Acak

##### 3.1.5. Hitung Jarak Data Ke Centroid

Setelah data dan algoritma K-Means dimasukkan ke Aplikasi Rapidminner maka proses data dengan mengklik tombol segitiga yang berwarna biru di pojok kiri untuk menjalankan algoritma sehingga dapat melihat jarak data ke Centroid. Dapat dilihat sebagai berikut.



Gambar 3.6 : Jarak Data Ke Centroid



Gambar 3.4 : Garfik Scatter 3D

$$d(x,y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2}$$

Sumber: (Suntoro, 2019)

Data ke-1

$$\begin{aligned} C0 &= \sqrt{(1 - 1.639)^2 + (1 - 1.833)^2} \\ &= \sqrt{(-0.639)^2 + (-0.833)^2} \\ &= \sqrt{0.408 + 0.693} \\ &= \sqrt{1.101} = 1.049 \end{aligned}$$

$$\begin{aligned} C1 &= \sqrt{(1 - 2.650)^2 + (1 - 6.813)^2} \\ &= \sqrt{(-1.560)^2 + (-5.813)^2} \\ &= \sqrt{2.433 + 33.79} \\ &= \sqrt{36.223} = 6.018 \end{aligned}$$

$$\begin{aligned}
C2 &= \sqrt{(1 - 4.644)^2 + (1 - 1.778)^2} \\
&= \sqrt{(-3.644)^2 + (-0.778)^2} \\
&= \sqrt{13.278 + 0.605} \\
&= \sqrt{13.883} = 3.725
\end{aligned}$$

Data ke-2

$$\begin{aligned}
C0 &= \sqrt{(1 - 1.639)^2 + (1 - 1.833)^2} \\
&= \sqrt{(-0.639)^2 + (-0.833)^2} \\
&= \sqrt{0.408 + 0.693} \\
&= \sqrt{1.101} = 1.049
\end{aligned}$$

$$\begin{aligned}
C1 &= \sqrt{(1 - 2.650)^2 + (1 - 6.813)^2} \\
&= \sqrt{(-1.560)^2 + (-5.813)^2} \\
&= \sqrt{2.433 + 33.79} \\
&= \sqrt{36.223} = 6.018
\end{aligned}$$

$$\begin{aligned}
C2 &= \sqrt{(1 - 4.644)^2 + (1 - 1.778)^2} \\
&= \sqrt{(-3.644)^2 + (-0.778)^2} \\
&= \sqrt{13.278 + 0.605} \\
&= \sqrt{13.883} = 3.725
\end{aligned}$$

Data ke-3

$$\begin{aligned}
C0 &= \sqrt{(1 - 1.639)^2 + (2 - 1.833)^2} \\
&= \sqrt{(-0.639)^2 + (-0.167)^2} \\
&= \sqrt{0.408 + 0.194} \\
&= \sqrt{0.602} = 0.775
\end{aligned}$$

$$\begin{aligned}
C1 &= \sqrt{(1 - 2.650)^2 + (2 - 6.813)^2} \\
&= \sqrt{(-1.560)^2 + (-4.813)^2} \\
&= \sqrt{2.433 + 23.164} \\
&= \sqrt{25.597} = 5.059
\end{aligned}$$

$$\begin{aligned}
C2 &= \sqrt{(1 - 4.644)^2 + (2 - 1.778)^2} \\
&= \sqrt{(-3.644)^2 + (-0.222)^2} \\
&= \sqrt{13.278 + 0.049} \\
&= \sqrt{13.327} = 3.650
\end{aligned}$$

Data Ke-4

$$\begin{aligned}
C0 &= \sqrt{(1 - 1.639)^2 + (1 - 1.833)^2} \\
&= \sqrt{(-0.639)^2 + (-0.833)^2} \\
&= \sqrt{0.408 + 0.693} \\
&= \sqrt{1.101} = 1.049
\end{aligned}$$

$$\begin{aligned}
C1 &= \sqrt{(1 - 2.650)^2 + (1 - 6.813)^2} \\
&= \sqrt{(-1.560)^2 + (-5.813)^2} \\
&= \sqrt{2.433 + 33.79} \\
&= \sqrt{36.223} = 6.018
\end{aligned}$$

$$\begin{aligned}
C2 &= \sqrt{(1 - 4.644)^2 + (1 - 1.778)^2} \\
&= \sqrt{(-3.644)^2 + (-0.778)^2} \\
&= \sqrt{13.278 + 0.605} \\
&= \sqrt{13.883} = 3.725
\end{aligned}$$

Data Ke-5

$$\begin{aligned}
C0 &= \sqrt{(5 - 1.639)^2 + (7 - 1.833)^2} \\
&= \sqrt{(3.361)^2 + (5.167)^2} \\
&= \sqrt{11.296 + 26.697} \\
&= \sqrt{37.993} = 6.163
\end{aligned}$$

$$\begin{aligned}
C1 &= \sqrt{(5 - 2.650)^2 + (7 - 6.813)^2} \\
&= \sqrt{(2.44)^2 + (0.187)^2} \\
&= \sqrt{5.953 + 0.034} \\
&= \sqrt{5.987} = 2.446
\end{aligned}$$

$$\begin{aligned}
C2 &= \sqrt{(5 - 4.644)^2 + (7 - 1.778)^2} \\
&= \sqrt{(0.356)^2 + (5.222)^2} \\
&= \sqrt{0.126 + 27.269} \\
&= \sqrt{27.395} = 5.234
\end{aligned}$$

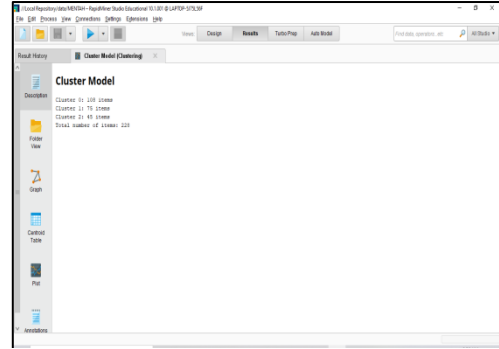
### 3.1.6. Perbarui Nilai Titik *Centroid*

Setelah melakukan percobaan, didapat titik centroid pada 3 cluster yang tidak mengalami perubahan.

### 3.1.7. Ulangi Langkah 3 dan 5 Sampai Nilai *Centroid*

#### Tidak Berubah

Apabila nilai titik pusat berubah maka akan di lakukan perulangan iterasi, tapi apabila nilai centroid tidak berubah maka alhasil perulangan di berhentikan sehingga telah mendapat masing-masing *cluster*.



Gambar 3. 7 : Hasil Cluster Pada Rapidminer

Berdasarkan perhitungan *rapidminer* maka diperoleh pola yang nantinya akan diimplementasikan di sistem. Menggunakan sebuah pola untuk mengelompokkan/*cluster* data berdasarkan hasil titik centroid adalah:

- Apabila  $C0 < C1$  dan  $C0 < C2$  yaitu *cluster 0* dengan keterangan tinggi
- Serta  $C1 < C0$  dan  $C1 < C2$  ialah *cluster 1* dengan keterangan sedang
- Dan  $C2 < C0$  dan  $C2 < C1$  adalah *cluster 2* dengan keterangan rendah

*Rapidminer* mendapatkan pola dengan metode *clustering k-means* maka diperoleh bahwa *cluster\_0* memiliki tingkat keluhan tinggi dengan jumlah 108, *cluster\_1* dengan keterangan tingkat keluhan sedang dengan jumlah 75 dan *cluster\_2* memiliki tingkat keluhan rendah dengan jumlah 45.

Setelah memperoleh pola dari *Rapid miner* maka penelitian ini menghasilkan sebuah *software* sistem *clustering* dengan algoritma K-Means yang akan disajikan pada Sistem data keluhan PT.PLN Persero Kota Pagar Alam Menggunakan Metode Clustering K-Means. Untuk memulai menggunakan aplikasi ini, dengan mengaktifkan Google Chrome.

## 3.2 Pembahasan

### 3.2.1. Pengujian *Black Box Testing*

Pengujian yang dilakukan dalam sistem ini menggunakan metode pengujian *black box* untuk memastikan bahwa sistem yang dihasilkan bebas dari kesalahan atau kegagalan dari keseluruhan sistem serta sesuai dengan apa yang di rancang.

#### 3.2.1.1. Rekapitulasi Hasil Uji Alpha

berdasarkan penilaian para sang ahli, pengujian database memperoleh nilai rata-rata 3,4, lalu antarmuka memperoleh nilai rata-rata 4, kemudian pengujian fungsionalitas menerima nilai rata-rata 4,6, serta pengujian algoritma memperoleh nilai rata-rata 4,7. pada beberapa tes sebelumnya, diperoleh nilai rata-rata 4. Berikut adalah hasil uji alpha.

hasil yang di terapkan di sistem yang di bangun dengan jumlah *cluster* yang terdiri dari 3 yaitu *cluster* 0 memiliki nilai *cluster* tertinggi, *cluster* 1 memiliki nilai *cluster* sedang dan *cluster* 2 memiliki nilai *cluster* rendah.

**Tabel 3. 1 : Hasil Uji Alpha**

| No    | Pengujian      | Rata-rata |
|-------|----------------|-----------|
| 1     | Database       | 3,4       |
| 2     | Antarmuka      | 4         |
| 3     | Fungsionalitas | 4,6       |
| 4     | Algoritma      | 4,7       |
| Total |                | 4,2       |

3. Pengujian *black box testing* berdasarkan hasil kuisisioner yang telah disebar oleh peneliti dan telah diisi oleh pakar berupa pengujian *alpha*, yang menghasilkan nilai kelayakan 84%.

Hal itu menunjukkan bahwa Penerapan Metode *Clustering K-Means* untuk menentukan data keluhan pelanggan agar di implementasikan.

### 3.2.1.2. Rekapitulasi Pengujian Alpha

Perhitungan dengan menggunakan

Analisis deskriptif seperti berikut:

$$\text{Persentase kelayakan } \alpha = \frac{\text{Skor yang diobservasi}}{\text{skor yang diharapkan}} \times 100\% = \frac{4,2}{5} \times 100\% = 84\%$$

Hasil perhitungan yang diperoleh digunakan dalam menentukan kelayakan media. Adapun kategori kelayakan adalah sebagai berikut:

**Tabel 3. 2 : Kategori Kelayakan**

| Kategori          | Persentase |
|-------------------|------------|
| Sangat Setuju(SS) | 81% - 100% |
| Setuju(S)         | 61% - 80%  |
| Ragu-ragu(RR)     | 41% - 60%  |
| Kurang Setuju(KS) | 21% - 40%  |
| Tdak Setuju(TS)   | 0% - 20%   |

Berdasarkan kuisisioner yang telah disebar oleh peneliti dan telah diisi oleh para pakar, dimana pengujian *database* mendapatkan nilai dengan skor rata-rata 3,4, kemudian pengujian algoritma mendapatkan nilai rata-rata 4,7, serta pengujian Antarmuka memperoleh nilai rata-rata 4 dan pengujian fungsionalitas dengan rata-rata 4,6. Dalam 4 tes tersebut, nilai rata-ratanya mendapat 4 sehingga mendapat nilai kelayakan 84%. Oleh sebab itu diperoleh bahwa sistem ini layak untuk digunakan atau tidak.

## 4. KESIMPULAN

Hasil yang didapat dari penelitian ini sebagai berikut:

1. Peneliti telah membuat sebuah sistem Penerapan Metode *Clustering K-Means* Untuk Menentukan Data Keluhan Pelanggan PT.PLN Persero Pagar Alam.
2. Hasil dari proses *clustering* data pengelompokan data Keluhan Pelanggan PT.PLN Persero dengan menggunakan apk *rapidminer* harus sama dengan

## 5. DAFTAR PUSTAKA

- Dahlan, A., Prasetyo, M., Erliana, C. I., Rahardja, U., & Karim, A. (2020). Sistem Informasi Pelayanan Dan Keluhan Pelanggan Di PT.PLN. In Sefa Bumi Persada. [https://repository.unimal.ac.id/5594/1/BUKU DAHLAN-MUHAJIR.pdf](https://repository.unimal.ac.id/5594/1/BUKU%20DAHLAN-MUHAJIR.pdf)
- Kasus, S., Keluhan, D., Pt, P., & Persero, P. L. N. (2022). PENERAPAN ALGORITMA K - MEANS CLUSTERING. 6(2), 327–340. <https://doi.org/10.52362/jisamar.v6i2.761>
- Hartanti, D., Mining, D., K-means, A., & Mining, D. (2015). Model Clustering Menggunakan Algoritma K-Means Pada Data Keluhan Pelanggan Pt . Pln ( Studi Kasus : Pt . Pln ( Persero ) Distribusi Jakarta Dan Tangerang ). Model Clustering Menggunakan Algoritma K-Means Pada Data Keluhan Pelanggan Pt . Pln ( Studi Kasus : Pt . Pln ( Persero ) Distribusi Jakarta Dan, 4, 119.
- Rohmatullah, F. M. (2018). Sistem Klasifikasi Keluhan Pelanggan Pt Pln Semarang Menggunakan Algoritma Naïve Bayes.
- Metisen, B. M., & Sari, H. L. (2015). Analisis clustering menggunakan metode K-Means dalam pengelompokkan penjualan produk pada Swalayan Fadhila. Jurnal Media Infotama, 11(2), 110–118.
- Suntoro, J. (2019). Data Mining Algoritma Dan Implementasi Dengan Pemrograman Php. Pt. Elex Media Komputindo.