

**YAYASAN LEMBAGA PENDIDIKAN ISLAM DAERAH RIAU
UNIVERSITAS ISLAM RIAU
FAKULTAS TEKNIK**

**IMPLEMENTASI DATA MINING K-MEANS *CLUSTERING*
TUNGGAKAN REKENING LISTRIK PASCABAYAR
(Studi Kasus : PT PLN Persero ULP Tualang)**

SKRIPSI

Diajukan Untuk Memenuhi Salah Satu Syarat
Penyusunan Skripsi Pada Fakultas Teknik
Universitas Islam Riau Pekanbaru

BALQIS MAIPASHA
173510557

**PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS TEKNIK
UNIVERSITAS ISLAM RIAU
PEKANBARU**

2022

LEMBAR PENGESAHAN PEMBIMBING SKRIPSI

Nama : Balqis Maipasha
NPM : 173510557
Fakultas : Teknik
Program Studi : Teknik Informatika
Jenjang Pendidikan : Strata Satu (S1)
Judul Skripsi : Implementasi Data Mining K-Means Clustering
Tunggakan Rekening Listrik Pascabayar (Studi Kasus :
PT. PLN (Persero) ULP Tualang).

Format sistematika dan pembahasan materi pada masing-masing bab dan sub bab dalam skripsi ini telah dipelajari dan dinilai relatif telah memenuhi ketentuan-ketentuan dan kriteria - kriteria dalam metode penulisan ilmiah. Oleh karena itu, skripsi ini dinilai layak dapat disetujui untuk disidangkan dalam ujian komprehensif.

Pekanbaru, 31 Januari 2022

Disahkan Oleh

Ketua Prodi Teknik Informatika

Dosen Pembimbing

Dr. APRI SISWANTO, S.Kom., M.Kom

AUSE LABELLAPANSA, S.T., M.Cs., M.Kom

LEMBAR PENGESAHAN TIM PENGUJI UJIAN SKRIPSI

Nama : Balqis Maipasha
NPM : 173510557
Fakultas : Teknik
Program Studi : Teknik Informatika
Jenjang Pendidikan : Strata Satu (S1)
Judul Skripsi : Implementasi Data Mining K-Means Clustering
Tunggakan Rekening Listrik Pascabayar (Studi
Kasus : PT PLN Persero ULP Tualang)

Skripsi ini secara keseluruhan dinilai telah memenuhi ketentuan-ketentuan dan kaidah-kaidah dalam penulisan penelitian ilmiah serta telah diuji dan dapat dipertahankan dihadapan tim penguji. Oleh karena itu, Tim Penguji Ujian Skripsi Fakultas Teknik Universitas Islam Riau menyatakan bahwa mahasiswa yang bersangkutan dinyatakan Telah Lulus Mengikuti Ujian Komprehensif Pada Tanggal 8 Februari 2022 dan disetujui serta diterima untuk memenuhi salah satu syarat guna memperoleh gelar Sarjana Strata Satu Bidang Ilmu Teknik Informatika.

Pekanbaru, 8 Februari 2021

Tim Penguji

1. Nesi Syafitri, S.Kom., M.Cs.

Sebagai Tim Penguji

(.....)
(.....)

2. Ir. Des Suryani, M.Sc

Sebagai Tim Penguji

Disahkan Oleh :

Ketua Prodi Teknik Informatika

Dosen Pembimbing



Dr. Apri Siswanto, S.Kom., M.Kom.



Ause Labellapansa, ST., M.Cs, M.Kom

LEMBAR PERNYATAAN BEBAS PLAGIARISME

Saya yang bertanda tangan dibawah ini :

Nama : Balqis Maipasha

NPM : 173510557

Fakultas : Teknik

Program Studi : Teknik Informatika

Jenjang Pendidikan : Strata-1 (S1)

Dengan ini menyatakan dengan sesungguhnya bahwa skripsi yang saya tulis adalah benar dan asli hasil dari penelitian yang telah saya lakukan dengan judul **Implementasi Data Mining K-Means Clustering Tunggalan Rekening Listrik Pascabayar (Studi Kasus : PT. PLN (Persero) ULP Tualang).**

Apabila dikemudian hari ada yang merasa dirugikan dan atau menuntut karena penelitian ini menggunakan sebagian hasil tulisan atau karya orang lain tanpa mencantumkan nama penulis yang bersangkutan, atau terbukti karya ilmiah ini **bukan** karya saya sendiri atau **plagiat** hasil karya orang lain, maka saya bersedia menerima sanksi sesuai dengan peraturan perundangan yang berlaku.

Demikian surat pernyataan ini saya buat dengan sesungguhnya untuk dapat digunakan sebagaimana mestinya.

Pekanbaru, 31 Januari 2022

Yang membuat pernyataan,



BALQIS MAIPASHA

KATA PENGANTAR

بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ

Assalaamu'alaikum Wr.Wb.

Segala puji bagi Allah SWT yang selalu memberikan rahmat dan hidayah-Nya serta nikmat yang tak terhingga, sehingga penulis dapat menyelesaikan proposal skripsi ini dengan judul **“IMPLEMENTASI DATA MINING K-MEANS CLUSTERING TUNGGAKAN REKENING LISTRIK PASCABAYAR (STUDI KASUS : PT PLN PERSERO ULP TUALANG)”** sebagai salah satu syarat untuk penyusunan laporan skripsi pada Fakultas Teknik Prodi Teknik Informatika Universitas Islam Riau.

Dalam penyusunan proposal ini, penulis sadar bahwa tanpa bantuan dan bimbingan berbagai pihak lain maka skripsi ini sulit untuk terwujud. Untuk itu dalam kesempatan ini penulis ingin menyampaikan ucapan terima kasih yang sebesar-besarnya kepada:

1. Bapak Dr. Eng. Muslim, ST., MT., selaku Dekan Fakultas Teknik Universitas Islam Riau.
2. Bapak Apri Siswanto ST., M.Kom, selaku Ketua Program Studi Teknik Informatika.
3. Ibu Ana Yulianti, ST., M.Kom, selaku Sekretaris Program Studi Teknik Informatika dan Kak Nur'ainaani selaku Admin Prodi TI yang telah ikhlas dan

sabar memberikan semangat dan dukungan kepada Penulis.

4. Dosen Penasehat Akademik, Dosen Pembimbing dan sekaligus Kepala Biro SIMFOKOM, Ibu Ause Labellapansa, ST., M.Cs., M.Kom. yang telah ikhlas dan sabar memberikan bimbingan dan arahan kepada Penulis disela-sela kesibukan beliau.
5. Kepada segenap Dosen Teknik Informatika, Universitas Islam Riau yang telah memberikan ilmu, pendidikan, dan pengetahuan kepada penulis selama duduk dibangku kuliah.
6. Kepada orang tua tercinta Ayah Zulkarnain dan Ibu Sumiati serta adik-adik yang telah banyak berkorban, mendo'akan, menyemangati dan mengisi dunia penulis dengan penuh kebahagiaan.
7. Kepada *The One and Only*, Suci Rahayu Maily Candra, S.Pd yang menemani penulis selama menempuh pendidikan dan selalu sabar dalam menghadapi penulis.
8. Kepada teman kos yaitu Kiting, Ncup dan Mpok Nur yang memberikan masukan dan arahan kepada penulis.
9. Kepada teman-teman Asisten Laboratorium TI dan seluruh teman-teman kelas C TI 2017.
10. Kepada Mas Adi Mustofa, ST yang telah mengajarkan dan membantu penulis.

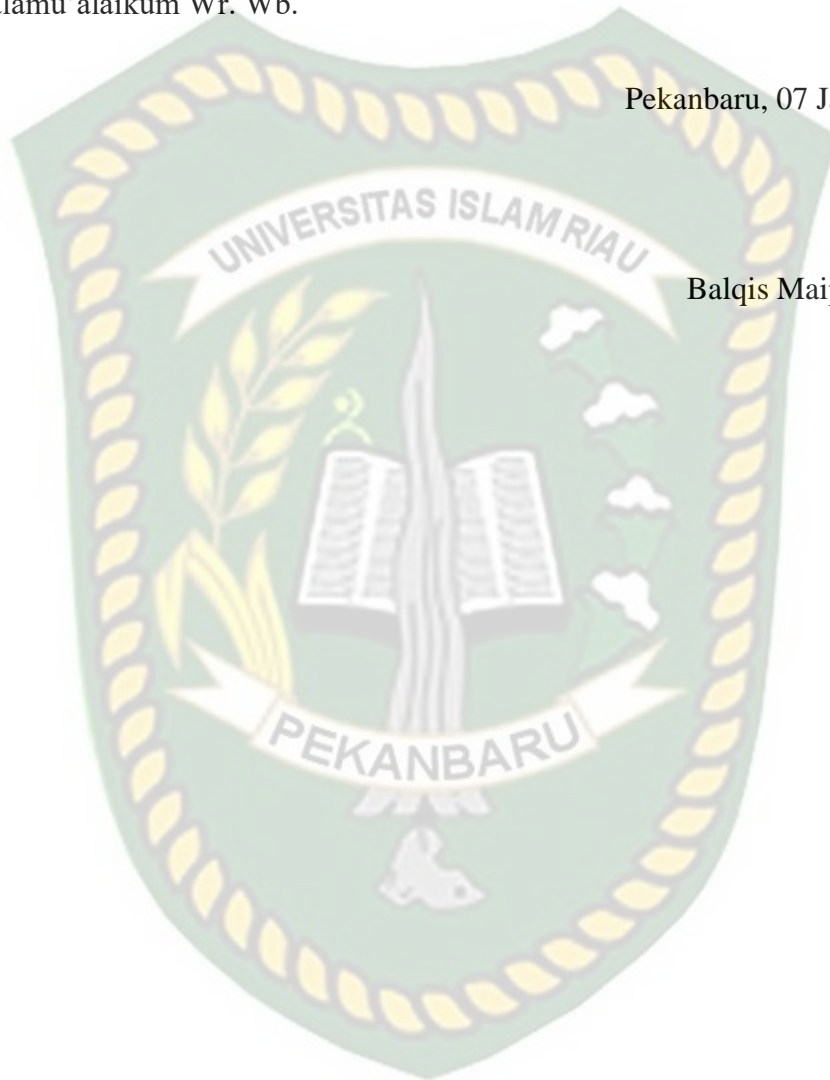
Penulis menyadari bahwa laporan skripsi ini masih jauh dari kesempurnaan oleh karena itu kritik dan saran dari pembaca sangat diharapkan untuk menyempurnakan skripsi ini.

Akhir kata semoga skripsi ini dapat menambah ilmu pengetahuan dan bermanfaat bagi semua pihak yang membacanya.

Wassalamu'alaikum Wr. Wb.

Pekanbaru, 07 Januari 2022

Balqis Maipasha



IMPLEMENTASI DATA MINING K-MEANS CLUSTERING TUNGGAKAN REKENING LISTRIK PASCABAYAR (STUDI KASUS : PT PLN PERSERO ULP TUALANG)

Balqis Maipasha

Mahasiswa Program Studi Teknik Informatika, Universitas Islam Riau

Jalan Kaharuddin Nasution Km.11 No. 113 Perhentian Marpoyan Pekanbaru

28284

Email: balqismaipasha@student.uir.ac.id

ABSTRAK

Banyaknya masyarakat yang masih menggunakan listrik pascabayar terutama di daerah kecamatan Tualang sehingga besarnya tunggakan pascabayar menjadi permasalahan terutama di PT PLN Persero Tualang. Banyak faktor yang mempengaruhi besarnya tunggakan seperti kurangnya kesadaran masyarakat, kecilnya pendapatan masyarakat saat pandemi, besarnya pemakaian listrik. Salah satu cara dalam menangani permasalahan tersebut adalah dengan melakukan tahap evaluasi terhadap masyarakat yang menunggak. Dengan Implementasi Data Mining K-Means *Clustering* ini diharapkan mampu membantu petugas PLN dalam mengambil keputusan yang tepat dalam memberikan tindakan serta sanksi dengan metode K-Means *Clustering*. Berdasarkan hasil analisa dan perancangan data dikelompokkan berdasarkan tarif yaitu tarif rumah tangga, tarif sosial dan tarif bisnis. Sistem mampu mengelompokkan data menjadi 3 klaster yaitu ditindaklanjuti, dalam monitoring dan aman dengan akurasi yaitu tarif rumah tangga sebesar 83%, tarif sosial 88% dan tarif bisnis 95% sehingga dapat disimpulkan bahwa sistem ini dapat diimplementasikan.

Kata Kunci: Klaster, k-means *clustering*, pascabayar, PT PLN, tunggakan Listrik

**IMPLEMENTATION OF DATA MINING K-MEANS
CLUSTERING POSTPAID ELECTRICITY ACCOUNTS
ARREARS
(CASE STUDY: PT PLN PERSERO ULP TUALANG)**

Balqis Maipasha

Student of Informatics Engineering Department, Universitas Islam Riau

Kaharuddin Nasution Street Km.11 No. 113, Marpoyan, Pekanbaru

28284

Email: balqismaipasha@student.uir.ac.id

ABSTRACT

The large number of people who still use postpaid electricity, especially in the Tualang sub-district, so that the amount of postpaid arrears is a problem, especially at PT PLN Persero Tualang. Many factors affect the amount of arrears, such as the lack of public awareness, the small income of the people during the pandemic, the amount of electricity consumption. One way to deal with these problems is to conduct an evaluation stage of the people who are in arrears. With the implementation of K-Means Clustering Data Mining, it is expected to be able to assist PLN officers in making the right decisions in providing actions and sanctions with the K-Means Clustering method. Based on the results of the analysis and design of the data, they are grouped based on tariffs, namely household tariffs, social tariffs and business tariffs. The system is able to classify data into 3 clusters, which are followed up, monitored and safe with accuracy, namely household rates of 83%, social rates of 88% and business tariffs of 95% so that it can be concluded that this system can be implemented.

Keywords: Cluster, k-means clustering, postpaid, PT PLN, electricity arrears.

DAFTAR ISI

KATA PENGANTAR	i
ABSTRAK	iv
ABSTRACT	v
DAFTAR ISI	vi
DAFTAR GAMBAR	viii
DAFTAR TABEL	ix
DAFTAR LAMPIRAN	xi
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Identifikasi Masalah	2
1.3 Batasan Masalah	2
1.4 Rumusan Masalah.....	3
1.5 Tujuan Penelitian	3
1.6 Manfaat Penelitian.....	3
BAB II LANDASAN TEORI	4
2.1 Tinjauan Pustaka.....	4
2.2 Dasar Teori	5
2.2.1 Tunggakan	5
2.2.2 Rekening Listrik	5
2.2.3 Data Mining	6
2.2.4 Tahapan Mining	6
2.2.5 <i>Clustering</i>	9
2.2.6 Metode K-Means	10
2.2.7 Tahapan – Tahapan <i>Clustering</i>	11
2.2.8 Perhitungan Manual K-Means.....	12
2.3 Dasar-Dasar Pemrograman.....	17
2.3.1 Basis Data	17
2.3.2 <i>Flowchart</i>	17
2.3.3 <i>Data Flow Diagram</i> (DFD).....	19
2.3.4 PHPStorm	19

2.3.5 <i>Laravel</i>	20
BAB III METODOLOGI PENELITIAN	21
3.1 Alat dan Bahan Penelitian.....	21
3.1.1 Alat Penelitian	21
3.1.2 Bahan Penelitian.....	22
3.2 Analisis Sistem yang Sedang Berjalan.....	22
3.3 Pengembangan Sistem	23
3.4.1 <i>Context Diagram</i>	24
3.4.2 <i>Hierarchy Chart</i>	24
3.4.3 Data Flow Diagram (DFD) Level 0.....	25
3.4 Rancangan Desain	26
3.5.1 Rancangan Desain <i>Output</i>	26
3.5.2 Rancangan Desain <i>Input</i>	27
3.5.3 Perancangan <i>Database</i>	30
3.5.4 Desain Antarmuka	32
3.5.5 Rancangan Logika Program	32
BAB IV HASIL DAN PEMBAHASAN.....	37
4.1 Pengujian <i>Black Box</i>	37
4.1.1 Pengujian <i>Form Training</i>	37
4.1.2 Pengujian <i>Proses K-Means</i>	40
4.1.3 Pengujian Proses Perhitungan Metode K-Means.....	41
4.1.4 Kesimpulan Pengujian <i>Black Box</i>	48
4.2 Pengujian <i>Whitebox</i>	48
4.2.1 Perhitungan Manual Metode K-Means.....	49
4.3 Implementasi Sistem	72
4.3.1 Kesimpulan Implementasi Sistem	74
BAB V PENUTUP.....	76
5.1 Kesimpulan	76
5.2 Saran	76
DAFTAR PUSTAKA	77

DAFTAR GAMBAR

Gambar 2. 1 Tahapan Mining.....	7
Gambar 3. 1 Analisis Sistem yang Berjalan	23
Gambar 3. 2 Pengembangan Sistem.....	23
Gambar 3. 3 Context Diagram.....	24
Gambar 3. 4 Hierarchy Chart	25
Gambar 3. 5 DFD Level 0.....	25
Gambar 3. 6 Output Data Training.....	26
Gambar 3. 7 Output Data Testing.....	26
Gambar 3. 8 Output Hasil Iterasi.....	27
Gambar 3. 9 Desain Input Login	28
Gambar 3. 10 Desain Input Training Manual.....	29
Gambar 3. 11 Desain Input Testing Manual.....	30
Gambar 3. 12 Desain Antarmuka	32
Gambar 3. 13 Flowchart Login.....	33
Gambar 3. 14 Flowchart Menu Sistem.....	34
Gambar 3. 15 Flowchart Input Training.....	35
Gambar 3. 16 Flowchart Algoritma K-Means.....	36
Gambar 4. 1 Pengujian Form Training.....	37
Gambar 4. 2 Tampilan Data Training.....	38
Gambar 4. 3 Tampilan Pesan Hapus	38
Gambar 4. 4 Pusat Cluster Awal	40
Gambar 4. 5 Hasil Clustering Tunggal.....	40
Gambar 4. 6 Pengujian Hasil Perhitungan Proses K-means.....	48
Gambar 4. 7 Hasil Perhitungan Sistem.....	55
Gambar 4. 8 Hasil Testing Sistem.....	56
Gambar 4. 9 Hasil Perhitungan Sistem Tarif Sosial.....	63
Gambar 4. 10 Hasil Testing Tarif Sosial	64
Gambar 4. 11 Hasil Perhitungan Sistem.....	71
Gambar 4. 12 Hasil Testing Tarif Rumah Tangga	72
Gambar 4. 13 Grafik Hasil Kuisisioner.....	73

DAFTAR TABEL

Tabel 2. 1 Contoh Perhitungan K-Means	13
Tabel 2. 2 Titik Centroid Awal.....	13
Tabel 2. 3 Matriks Perhitungan Jarak Iterasi Pertama.....	14
Tabel 2. 4 Pengelompokan Cluster Iterasi Pertama	15
Tabel 2. 5 Pusat Cluster Baru	15
Tabel 2. 6 Matriks Perhitungan Jarak Iterasi Ke-2.....	16
Tabel 2. 7 Perbandingan Iterasi	16
Tabel 2. 8 Simbol - Simbol Flowchart	18
Tabel 2. 9 Simbol DFD dan Fungsinya	19
Tabel 3. 1 Tabel User	30
Tabel 3. 2 Tabel Training	31
Tabel 3. 3 Tabel Data Testing	31
Tabel 4. 1 Kesimpulan Pengujian Form Training	39
Tabel 4. 2 Kesimpulan Pengujian Form Proses K-Means.....	41
Tabel 4. 3 Data Training Tarif Bisnis.....	49
Tabel 4. 4 Titik Pusat Cluster Awal	50
Tabel 4. 5 Perhitungan Jarak Iterasi-1	51
Tabel 4. 6 Pengelompokan Cluster Iterasi-1	52
Tabel 4. 7 Pusat Cluster Baru.....	53
Tabel 4. 8 Perhitungan Jarak Iterasi-2.....	53
Tabel 4. 9 Perbandingan Iterasi	54
Tabel 4. 10 Data Training Tarif Sosial.....	56
Tabel 4. 11 Pusat Cluster Awal	57
Tabel 4. 12 Perhitungan Jarak Iterasi-1	58
Tabel 4. 13 Pengelompokan Cluster Pertama.....	59
Tabel 4. 14 Pusat Cluster Baru.....	60
Tabel 4. 15 Perhitungan Jarak Iterasi-2.....	60
Tabel 4. 16 Perbandingan Iterasi	62
Tabel 4. 17 Data Training Tarif Rumah Tangga	64
Tabel 4. 18 Inisialisasi Pekerjaan	65

Tabel 4. 19 Pusat Cluster Awal	65
Tabel 4. 20 Perhitungan Jarak Iterasi-1	66
Tabel 4. 21 Pengelompokan Cluster Iterasi-1	68
Tabel 4. 22 Pusat Cluster Baru.....	69
Tabel 4. 23 Perhitungan Jarak Iterasi-2.....	69
Tabel 4. 24 Perbandingan Iterasi	70
Tabel 4. 25 Perhitungan Kuisisioner dengan Skala Likert.....	74



DAFTAR LAMPIRAN

- Lampiran 1** Kuisioner Implementasi Sistem
- Lampiran 2** SK Skripsi
- Lampiran 3** Kartu Bimbingan Skripsi
- Lampiran 4** Transkrip Nilai
- Lampiran 5** SK Penguji Komprehensif
- Lampiran 6** Berita Acara Komprehensif
- Lampiran 7** Surat Keterangan Bebas Plagiat



BAB I

PENDAHULUAN

1.1 Latar Belakang

Perusahaan Listrik Negara atau PT. PLN (Persero) merupakan suatu perusahaan yang bergerak dalam bidang pelayanan jasa energi listrik di Indonesia. Sebagai penyedia tenaga kelistrikan satu-satunya di Indonesia, PLN selalu berusaha melayani masyarakat Indonesia dalam memenuhi kebutuhan akan sumber daya listrik yang baik untuk pelanggannya.

PLN menyediakan layanan listrik pascabayar, yaitu pelanggan menggunakan energi listrik dulu dan membayar belakangan pada bulan berikutnya. Setiap bulan PLN harus mencatat meter, menghitung dan menerbitkan rekening yang harus dibayar pelanggan, melakukan penagihan kepada pelanggan yang menunggak atau tidak membayar rekening listrik setelah waktu tertentu.

Beberapa faktor yang mempengaruhi pelanggan pascabayar menunggak, seperti pencatatan meteran yang salah, minimnya penghasilan pokok perbulan pelanggan PLN, alasan klasik yang menghambat pelanggan untuk segera membayar, serta ketidaksadaran pelanggan mengenai pentingnya membayar listrik tepat waktu.

Oleh karena itu, pentingnya mengantisipasi pelanggan yang menunggak dengan melakukan penegasan wajib bayar tepat waktu. Akan tetapi kendalanya adalah besarnya *budget* yang dikeluarkan jika mendatangi semua pelanggan yang menunggak dan tidak efisien waktu yang akan digunakan.

1.2 Identifikasi Masalah

Berdasarkan uraian latar belakang diatas, maka dapat dibuat suatu identifikasi masalah yaitu :

1. Besarnya *budget* yang dikeluarkan untuk sosialisasi ke seluruh pelanggan yang menunggak.
2. Kurangnya kesadaran pelanggan dalam membayar tagihan tepat waktu.
3. Terhambatnya biaya operasional lainnya sehingga mengalokasikan dana tunggakan untuk keperluan operasional lain.
4. Belum adanya pengelompokan data tunggakan masing-masing tarif.

1.3 Batasan Masalah

Dengan adanya maksud dan tujuan dalam penelitian ini, maka penulis membuat batasan masalah yaitu :

1. Penelitian ini mengelompokkan data berdasarkan tarif dengan setiap tarif menggunakan atribut yang berbeda.
2. Penelitian ini menggunakan metode algoritma K-Means *clustering*.
3. Ruang lingkup *clustering* sistem hanya terbatas pada daerah kecamatan Tualang dengan 7 desa di kecamatan Tualang.
4. Sistem yang dibuat untuk membantu penelitian.

1.4 Rumusan Masalah

Adapun Rumusan Masalah dari penelitian ini adalah sebagai berikut:

1. Bagaimana implementasi dari data *mining* dengan menggunakan metode K-Means ?
2. Bagaimana cara membentuk *cluster* berdasarkan tarifnya masing-masing ?

1.5 Tujuan Penelitian

Adapun tujuan dari penelitian ini adalah :

1. Merancang dan menerapkan data *mining* dalam kluster data tunggakan rekening listrik pelanggan pascabayar.
2. Dengan dibangunnya sistem ini, pihak PLN dapat melakukan evaluasi, observasi dan sosialisasi terhadap pelanggan yang menunggak sehingga biaya operasional dalam mengatasi hal tersebut dapat diminimalisir.

1.6 Manfaat Penelitian

Adapun manfaat dari penelitian ini adalah sebagai berikut :

1. Menertibkan pelanggan yang menunggak agar segera membayar tagihannya sehingga pelanggan pascabayar menyadari betapa pentingnya membayar tagihan tepat waktu.
2. Meminimalisir biaya operasional dalam penagihan tunggakan rekening listrik pascabayar.

BAB II

LANDASAN TEORI

2.1 Tinjauan Pustaka

Dalam penelitian *clustering* K-Means, peneliti menggunakan beberapa kajian yang berhubungan dengan algoritma K-Means, dari penelitian yang telah dilakukan sebelumnya, yaitu:

Penelitian yang dilakukan oleh Ediman Manik 2018 dalam jurnal Data Mining Pengelompokan Pemakaian Listrik Terboros Menggunakan Metode *Clustering* (Studi Kasus PLN Area Binjai) adalah penelitian yang bertujuan untuk mengelompokkan pelanggan yang menggunakan listrik secara berlebihan sesuai dengan ketentuan batasan aturan penggunaan listrik setiap pelanggan.

Penelitian yang dilakukan oleh Rozi Kesuma Dinata, dkk 2020 dalam jurnal Analisis K-Means *Clustering* pada Data Sepeda Motor adalah penelitian yang bertujuan untuk mengelompokkan data sepeda motor menjadi 3 *cluster* yaitu murah, standar dan mahal yang digunakan untuk rekomendasi bagi pengguna dalam menentukan pemilihan sepeda motor yang diinginkan.

Penelitian yang dilakukan oleh Nurul Rohmawati, dkk 2015 dalam jurnal Implementasi Algoritma K-Means Dalam Pengklasteran Mahasiswa Pelamar Beasiswa adalah penelitian yang bertujuan untuk mengelompokkan data pelamar beasiswa Bantuan Belajar Mahasiswa (BBM) dalam 3 kelompok yaitu berhak menerima, dipertimbangkan dan tidak berhak menerima.

2.2 Dasar Teori

2.2.1 Tunggakan

Tunggakan adalah angsuran, pajak atau tagihan yang belum dibayar selama seseorang, lembaga ataupun badan usaha yang masih menggunakan jasa tersebut.

Pihak PLN memiliki peraturan yang sudah ditetapkan tentang keterlambatan dalam membayar tagihan listrik setiap bulannya, yaitu :

1. PLN menetapkan batasan akhir masa pembayaran tagihan tanggal 20 setiap bulannya.
2. Pelanggan pascabayar yang melampaui batasan akhir tagihan akan dikenakan biaya keterlambatan (BK).
3. Selain dikenakan BK, pelanggan juga dikenakan sanksi pemutusan.
4. Berlakunya BK hanya dibatasi sebanyak 3 (tiga) kali tarif BK, artinya pelanggan hanya bisa menunggak selama 3 bulan. Setiap tunggakan pertama, kedua dan ketiga memiliki sanksi masing-masing.

Dalam kondisi pandemi saat ini, kondisi ekonomi masyarakat sangat memprihatinkan. Sulitnya masyarakat untuk memenuhi kebutuhan rumah tangga termasuk arus listrik. Maka kemungkinan besar akan banyak terjadi penunggakan tagihan listrik di masa pandemi.

2.2.2 Rekening Listrik

Rekening listrik merupakan perhitungan biaya pemakaian listrik yang wajib dibayar oleh para pelanggan setiap bulannya. Hasil perhitungan jumlah

biaya pemakaian listrik ini akan berfungsi sebagai dokumen perusahaan dan sebagai bukti pembayaran bagi pelanggan.

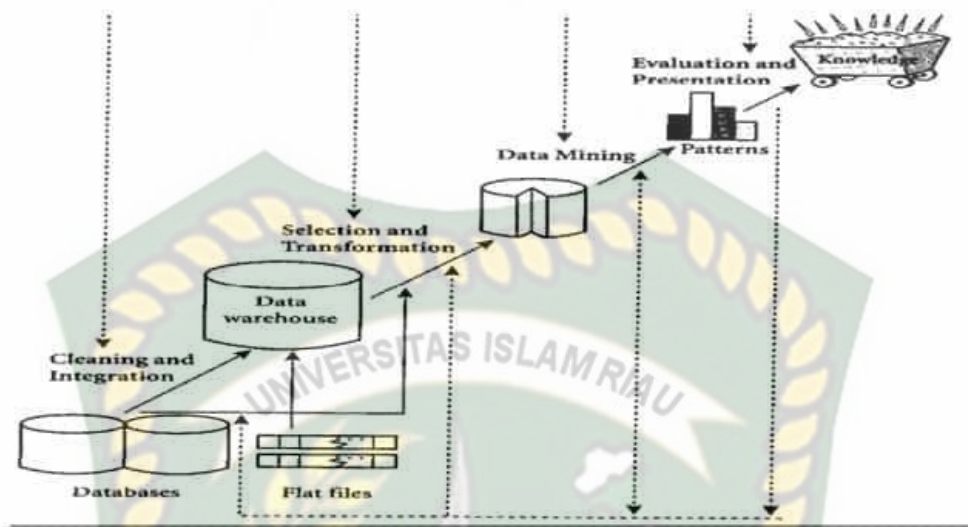
2.2.3 Data Mining

Menurut Vulandari (2017), “Data Mining adalah serangkaian proses untuk menggali nilai tambah berupa informasi yang selama ini tidak diketahui secara manual dari suatu basis data. Informasi yang dihasilkan diperoleh dengan cara mengekstraksi dan mengenali pola yang penting atau menarik dari data yang terdapat pada basis data. Data mining terutama digunakan untuk mencari pengetahuan yang terdapat dalam basis data yang besar sehingga sering disebut *Knowledge Discovery Databases (KDD)*”.

Menurut Suyanto (2019), “Data Mining merupakan gabungan sejumlah disiplin ilmu komputer yang didefinisikan sebagai proses penemuan pola-pola baru dari kumpulan-kumpulan data sangat besar, meliputi metode-metode yang merupakan irisan dari *artificial intelligence, machine learning, statistic*, dan *database systems*”.

2.2.4 Tahapan Mining

Ada beberapa tahapan dalam data mining. Diagram dibawah menggambarkan beberapa tahap/proses yang berlangsung dalam data mining. Fase awal dimulai dari data sumber dan berakhir dengan adanya informasi yang dihasilkan dari beberapa tahapan, yaitu :



Gambar 2. 1 Tahapan *Mining*

1. *Data Cleansing*

Data cleansing atau pembersihan data merupakan sebuah fase yang belum lengkap, masih terdapat error dan tidak konsisten. Dengan meminimalisir kemungkinan adanya data yang tidak sempurna agar tidak mempengaruhi performa dari tahapan data mining yang mengakibatkan berkurangnya jumlah dan kompleksitas data sehingga data yang tidak relevan tidak dapat digunakan untuk diproses ulang pengetahuannya (*discovery knowledge*).

2. *Data Integration*

Dalam tahapan ini dilakukan kombinasi atau penggabungan data yang diperoleh dari sumber atau file yang berulang (*multiple*) ke dalam *database* baru agar tidak terdapat data yang *multiple* sehingga data tidak menyimpang saat pengambilan hasil nantinya.

3. *Selection*

Tahap ini mengambil data yang sudah relevan terhadap analisis yang dapat diterima oleh kumpulan data lainnya. Karena data yang diperoleh dari sumber sering kali tidak semuanya dipakai, oleh karena itu hanya data yang sesuai dengan kasusnya untuk dianalisis yang akan diambil dari sumber data tersebut.

4. *Data Transformation*

Data yang telah terpilih akan ditransformasikan kedalam bentuk prosedur dengan cara melakukan normalisasi data dan agregasi untuk diproses dengan beberapa metode seperti *clustering* yang dapat menerima inputan data kategorikal dan data yang berupa numerik akan dibagi menjadi beberapa interval.

5. *Data Mining*

Tahapan ini paling penting karena menggunakan pola-pola yang sudah diekstrak untuk diimplementasikan sehingga dapat ditemukan informasi yang jelas.

6. *Pattern Evaluation*

Pattern evaluation atau evaluasi pola bertujuan untuk mengidentifikasi pola-pola yang menarik kedalam *knowledge based*. Pada tahap evaluasi ini, pola ataupun model yang didapatkan akan dianalisis apakah hipotesa yang ada sudah tercapai.

7. *Knowledge Representation*

Knowledge representation atau presentasi pengetahuan merupakan tahap terakhir dari data mining yang akan dipresentasikan terhadap user dengan menampilkan hasil visualisasi dari pengetahuan yang telah ditemukan agar dapat

dipahami user dengan memformulasikan keputusan atau hasil analisis yang didapat secara *user friendly*.

2.2.5 Clustering

Menurut Vulandari (2017), “*Clustering* adalah metode data mining dengan konsep *unsupervised learning*, karena tidak ada satu atributpun yang digunakan untuk memandu proses pembelajaran, jadi seluruh atribut input diperlakukan sama. Kebanyakan algoritma *clustering* membangun sebuah model melalui serangkaian pengulangan dan berhenti ketika model tersebut telah memusat atau berkumpul (batasan dari segmentasi ini telah stabil)”.

Clustering juga disebut sebagai *segmentation*. Metode ini digunakan untuk mengidentifikasi kelompok alami dari sebuah kasus yang didasarkan pada sebuah kelompok atribut, mengelompokkan data yang memiliki kemiripan atribut.

Pengklusteran berbeda dengan klasifikasi yaitu tidak adanya variabel target dalam pengklusteran. Pengklusteran tidak mencoba untuk melakukan klasifikasi, mengestimasi atau memprediksi nilai dari variabel target. Akan tetapi, algoritma pengklusteran mencoba untuk melakukan pembagian terhadap keseluruhan data menjadi kelompok-kelompok yang memiliki kemiripan (homogen), yang mana kemiripan *record* dalam satu kelompok akan bernilai maksimal sedangkan kemiripan dengan *record* dalam kelompok lain akan bernilai minimal.

2.2.6 Metode K-Means

K-Means merupakan salah satu algoritma yang bertujuan untuk membagi data menjadi beberapa kelompok. Algoritma ini menerima masukan berupa data tanpa label kelas. Komputer mampu mengelompokkan sendiri data-data yang menjadi masukannya tanpa mengetahui terlebih dulu target kelasnya. Masukan yang diterima adalah data atau objek dan k buah kelompok yang diinginkan. Pada setiap *cluster* terdapat titik pusat (*centroid*) yang mempresentasikan *cluster* tersebut. Secara sederhana, K-Means adalah algoritma *machine learning* yang digunakan untuk menyelesaikan permasalahan *clustering*.

Berdasarkan jurnal yang dikarang oleh Nurul Rohmawati, dkk yang mengemukakan algoritma K-Means sebagai berikut:

1. Menentukan k sebagai jumlah kluster yang ingin di bentuk.
2. Membangkitkan nilai *random* untuk pusat kluster awal (*centroid*) sebanyak k .
3. Menghitung jarak setiap data *input* terhadap masing-masing *centroid* menggunakan rumus jarak *Euclidean Distance* hingga ditemukan jarak yang paling dekat dari setiap data dengan *centroid*. Berikut adalah persamaan *Euclidean Distance*.

$$d(x_i, \mu_j) = \sqrt{\sum (x_i - \mu_j)^2} \dots \dots \dots (2.1)$$

Dimana :

x_i = data kriteria

μ_j = *centroid* pada kluster ke- j

4. Mengklasifikasikan setiap data berdasarkan kedekatannya dengan *centroid* (jarak terkecil).

5. Memperbaharui nilai *centroid*. Nilai *centroid* baru diperoleh dari rata-rata *cluster* yang bersangkutan dengan menggunakan rumus :

$$\mu_j(t+1) = \frac{1}{N_{sj}} \sum_{j \in S_j} x_j \dots \dots \dots (2.2)$$

Dimana :

$\mu_j(t+1)$ = *centroid* baru pada iterasi ke (t + 1)

N_{sj} = banyak data pada *cluster* S_j .

Melakukan perulangan dari langkah 2 hingga 5, sampai anggota tiap *cluster* tidak ada yang berubah.

Berikut ini *pseudocode* dari metode k-means *clustering* :

Pilih sejumlah k objek secara acak dari himpunan data D sebagai *centroid*

awal

Repeat

For semua objek di dalam D

Masukkan setiap objek yang bukan *centroid* ke klaster yang paling dekat di antara k klaster yang ada

End

Perbarui setiap *centroid* dengan menghitung rata-rata dari semua objek

Until tidak ada perubahan *centroid*.

2.2.7 Tahapan – Tahapan *Clustering*

Tahapan dilakukan agar dapat menyelesaikan masalah secara terstruktur.

Ada beberapa tahapan yang harus dilakukan dalam proses pengklasteran, yaitu :

1. Tahap Pengumpulan Data

Tahap pengumpulan data merupakan tahap yang mempelajari materi teori-teori penelitian yang berkaitan dengan identifikasi masalah. Proses tahap pengumpulan data ini mulai dari membaca studi literatur yang berkaitan dengan permasalahan penelitian, untuk menentukan atribut atau variabel yang digunakan dalam penelitian yang akan dilakukan. Kemudian masalah dalam kasus penelitian diidentifikasi untuk mendapatkan tujuan dari penelitian.

2. Tahap *Training*

Tahap *training* merupakan tahap yang digunakan untuk proses data latih sebelum dilakukan proses pengelompokan data tunggakan rekening listrik pascabayar menggunakan metode K-Means.

3. Tahap *Testing*

Tahap *testing* merupakan tahap yang digunakan untuk proses pengujian untuk melihat keberhasilan dari suatu sistem dalam membaca data uji (*testing*). Beberapa data *training* akan dipilih sebagai data *testing* yang akan diuji untuk melakukan proses klaster.

2.2.8 Perhitungan Manual K-Means

Perancangan pada analisa data ini terdiri dari data tunggakan rekening listrik pada bulan April 2021 di PLN ULP Tualang.

Pada pengelompokkan ini, ada 3 tarif yang terdaftar dalam tunggakan bulan April 2021 yaitu tarif rumah tangga, tarif sosial dan tarif bisnis. Setiap data akan dikelompokkan berdasarkan tarifnya.

Percobaan dilakukan dengan menggunakan parameter-parameter berikut ini:

Jumlah *Cluster* : 3

Jumlah data : 6

Jumlah atribut : 4 (Daya, Rupiah Tagihan, kWh, Jam Nyala)

Keterangan :

kWh : Pemakaian energi listrik dalam satuan kWh dalam waktu satu bulan.

Jam Nyala : Rasio dari pemakaian kWh dalam satu bulan dibagi dengan daya (kVA).

Tabel 2. 1 Contoh Perhitungan K-Means

No.	Nama	Daya	Tagihan	kWh	Jam Nyala
1	DESMAN	900	Rp105,980.00	192	213
2	YUSNIMAR.II	5500	Rp1,182,170.00	977	178
3	DARYULISMAN	1300	Rp87,133.00	82	63
4	SAHAR	1300	Rp218,896.00	206	158
5	NASRUL	2200	Rp72,600.00	32	15
6	Ely Ningsih	1300	Rp518,549.00	488	375

Iterasi Ke-1 :

- a. Menentukan pusat awal *Cluster*

Nilai Random awal :

Tabel 2. 2 Titik *Centroid* Awal

	Daya	Tagihan	kWh	Jam Nyala
C1	5500	Rp1,182,170.00	977	178
C2	2200	Rp72,600.00	32	15
C3	1300	Rp87,133.00	82	63

b. Perhitungan jarak pusat *cluster*

Untuk mengukur jarak antara data dengan pusat *cluster* digunakan *Euclidean distance* dengan rumus sebagai berikut :

Berikut contoh perhitungan jarak dari data ke-1 terhadap pusat *cluster* :

$$d(x_1, c_1) = \sqrt{(900 - 5500)^2 + (105980 - 1182170)^2 + (192 - 977)^2 + (213 - 178)^2} = 1076200$$

$$d(x_1, c_2) = \sqrt{(900 - 2200)^2 + (105980 - 72600)^2 + (192 - 32)^2 + (213 - 15)^2} = 33406$$

$$d(x_1, c_3) = \sqrt{(900 - 1300)^2 + (105980 - 87133)^2 + (192 - 82)^2 + (213 - 63)^2} = 18852$$

Dan seterusnya dilanjutkan untuk data ke 2, 3,...n. Kemudian akan didapatkan matrik jarak untuk iterasi ke-1.

Tabel 2. 3 Matriks Perhitungan Jarak Iterasi Pertama

NO	NAMA	C1	C2	C3	Cluster
1	DESMAN	1076200	26129	175063	C3
2	YUSNIMAR.II	0	1109575	1095045	C1
3	DARYULISMAN	1095045	14561	0	C2
4	SAHAR	963283	146299	131763	C3
5	NASRUL	1109575	0	14561	C2
6	Ely Ningsih	663634	445950	431416	C3

Setiap kolom pada matrik menunjukkan nilai jarak data terhadap pusat *cluster*. Baris pertama pada matrik menunjukkan nilai jarak data terhadap titik pusat *cluster* pertama, dan baris kedua pada matriks menunjukkan nilai jarak data terhadap titik pusat *cluster* kedua begitu juga dengan baris ketiga.

c. Pengelompokan data

Jarak dari hasil perhitungan yang sudah dilakukan, akan dibandingkan dan dipilih jarak terdekat atau nilai terkecil dari ketiga hasil perhitungan jarak setiap *cluster*, jarak ini menunjukkan bahwa data tersebut berada dalam satu kelompok dengan *cluster* terdekat.

Tabel 2. 4 Pengelompokan *Cluster* Iterasi Pertama

NO	C1	C2	C3
1			*
2	*		
3		*	
4			*
5		*	
6			*

d. Perhitungan titik *centroid* baru

Hitung titik *centroid* baru berdasarkan data anggota kelompok tiap-tiap *cluster* sesuai dengan rumus titik *centroid*. Sehingga didapatkan sebagai berikut :

$$C1 = \left(\frac{5500}{1}, \frac{1182170}{1}, \frac{977}{1}, \frac{178}{1} \right) = (5500; 1182170; 977; 178)$$

$$C2 = \left(\frac{1300+2200}{2}, \frac{87133+72600}{2}, \frac{82+32}{2}, \frac{63+15}{2} \right) \\ = (1750; 79866,5; 57; 39)$$

$$C3 = (1166,6; 281043; 295,3; 248,6)$$

Tabel 2. 5 Pusat *Cluster* Baru

	DAYA	TAGIHAN	PEMKWH	JAMNYALA
C1	5500	1182170	977	178
C2	1750	79866	57	39
C3	1166,6	281043	295,3	248,6

Iterasi ke-2 :

Ulangi langkah kedua (b) hingga data tidak mengalami perubahan. Berikut hasil perhitungan jarak iterasi kedua :

Tabel 2. 6 Matriks Perhitungan Jarak Iterasi Ke-2

NO	NAMA	C1	C2	C3	Cluster
1	DESMAN	1076200	26129	175063	C2
2	YUSNIMAR.II	1102311	901138		C1
3	DARYULISMAN	1095045	7281	193910	C2
4	SAHAR	963283	139031	62147	C3
5	NASRUL	1109575	7280	208446	C2
6	Ely Ningsih	663634	438684	237506	C3

Langkah berikutnya, melakukan perbandingan perbandingan hasil *cluster* pada iterasi pertama dengan kedua. Jika nilai anggota *cluster* memiliki nilai yang konvergen/sama maka proses iterasi berhenti. Iterasi dihentikan dan hasil *cluster* tersebut merupakan hasil *cluster* yang sudah valid.

Tabel 2. 7 Perbandingan Iterasi

NO	Iterasi 1	Iterasi 2
1	C3	C2
2	C1	C1
3	C2	C2
4	C3	C3
5	C2	C2
6	C3	C3

Kesimpulannya, jika nilai anggota *cluster* sebelumnya memiliki nilai yang konvergen dengan nilai anggota *cluster* setelahnya maka proses iterasi dihentikan. Jika nilai anggota *cluster* belum konvergen maka ulangi langkah sebelumnya (d) untuk menghitung pusat *cluster* baru.

2.3 Dasar-Dasar Pemrograman

2.3.1 Basis Data


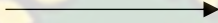


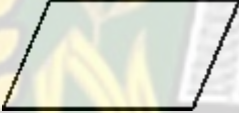




Menurut Solichin (2010), “Basis data atau *database* adalah kumpulan data yang terintegrasi dengan baik dan secara sistematis disimpan di dalam komputer agar memperoleh informasi yang jelas dan aman. Data yang terdiri dari kumpulan file-file yang saling berelasi dan dihubungkan dengan kunci atau *primary key* dan *foreign key*”.

Database dirancang agar memudahkan *user* untuk mendapatkan informasi sesuai kebutuhan sistemnya. Untuk memudahkan user dalam mengelola *database* maka diperlukan *software* DBMS (*Database Management System*). DBMS merupakan perangkat lunak yang mendukung pemrosesan data seperti menyimpan data, memelihara, mengontrol dan mengakses *database* secara praktis dan efisien. Beberapa contoh DBMS yang populer seperti *phpMyAdmin*, *Oracle*, *MS Access*, *SQLyog*, dan lainnya.

2.3.2 Flowchart

Flowchart adalah bagan-bagan yang mempunyai arus yang menggambarkan langkah-langkah penyelesaian suatu masalah. *Flowchart* merupakan cara penyajian dari suatu algoritma. Simbol *flowchart* dan fungsinya dapat dilihat pada tabel berikut (Ladjamudin, 2006:265) :

Tabel 2. 8 Simbol - Simbol *Flowchart*

No	Simbol	Nama	Fungsi
1		Terminator	Memulai alur/ alur berhenti
2		Flow Line	Arah penghubung
3		Preparation	Memberi variabel dan nilai awal
4		Process	Proses perhitungan data / array
5		Input/output data	Deskripsi data (read), penulisan (write)
6		Predefined Process	Fungsi prosedur / sub alur program
7		Decision	Pernyataan pilihan dengan menyeleksi proses menghasilkan nilai benar atau salah
8		On Page Connector	Menghubungkan flowchart dalam halaman yang sama
9		Off page Connector	Menghubungkan halaman flowchart yang terpisah

2.3.3 Data Flow Diagram (DFD)

Menurut Sukrianto (2017) menjelaskan bahwa, “DFD merupakan gambaran sistem secara logika yang tidak tergantung pada perangkat keras, lunak, struktur data dan organisasi file”. Keuntungan DFD adalah untuk memudahkan pemakai yang kurang menguasai bidang komputer untuk mengerti sistem yang akan dikerjakan atau dikembangkan. Berikut simbol-simbol dari DFD :

Tabel 2. 9 Simbol DFD dan Fungsinya

Simbol	Nama	Fungsi
	Simbol entitas eksternal	Objek yang berinteraksi dengan sistem tetapi diluar sistem
	Simbol proses	Melakukan transformasi data dalam proses
	Simbol penyimpanan <i>data</i>	Tempat data diorder untuk diproses
	Simbol arus <i>data</i>	Aliran data dengan arah khusus dari sumber ke tujuan

2.3.4 PHPStorm

PhpStorm merupakan perangkat lunak berupa IDE yang kompatibel dengan bahasa pemrograman PHP, *Javascript* dan HTML yang berbasis *Intellij IDEA*. PhpStorm dikembangkan oleh JetBrains yang memfokuskan pada versi PHP. PhpStorm bukan perangkat lunak yang *open source*, namun JetBrains menawarkan lisensi gratis untuk institud akademik.

2.3.5 *Laravel*

Laravel merupakan *framework open source* yang digunakan untuk bahasa pemrograman PHP dengan menggunakan konsep MVC (*Model View Controller*) yang dapat membantu kinerja sistem agar lebih efektif dan efisien dari pada menggunakan bahasa pemrograman yang statis. *Laravel* juga memiliki *libraries* yang memudahkan *user* dalam membangun sistemnya.



BAB III

METODOLOGI PENELITIAN

3.1 Alat dan Bahan Penelitian

3.1.1 Alat Penelitian

Pada penelitian ini penulis menggunakan alat dan bahan sebagai pendukung Implementasi data mining dengan metode K-Means *Clustering* pada PT PLN Persero ULP Tualang. Berikut kebutuhan spesifikasi pada penelitian ini :

3.1.1.1 Spesifikasi Kebutuhan *Hardware*

Struktur Perangkat Keras (*Hardware*) haruslah memenuhi spesifikasi kebutuhan dari sistem yang digunakan, berikut spesifikasi *Hardware* yang digunakan :

1. *Processor* : Intel(R) Core(TM) i3-5005U
2. *RAM* : 2,00 GB
3. *Hardisk* : 500 GB
4. *System Type* : 64-bit *Operating System*

3.1.1.2 Spesifikasi Kebutuhan *Software*

Perangkat Lunak (*Software*) yang digunakan dalam pembuatan sistem implementasi data mining tunggakan rekening listrik adalah :

1. Sistem Operasi : *Microsoft Windows 10 Enterprise*
2. Bahasa Pemrograman : HTML, PHP
3. *Database Management System* : *SQLYog*
4. *Web Browser* : *Google Chrome*

3.1.2 Bahan Penelitian

3.1.2.1 Jenis Data Penelitian

Adapun jenis data penelitian yang digunakan dalam penelitian ini menggunakan data primer. Data sekunder berasal dari data yang tersedia dari PT PLN Persero ULP Tualang.

3.1.2.2 Metode Pengumpulan Data

Dalam metode pengumpulan data yang diperlukan dalam membangun sistem ini harus menggunakan data yang akurat agar kasus penelitian ini sesuai dengan hasil yang akan dikerjakan penulis. Berikut metode pengumpulan data dalam penelitian ini :

1. Wawancara

Wawancara dilakukan untuk mendapatkan data secara langsung dari sumbernya yaitu PT PLN Persero ULP Tualang.

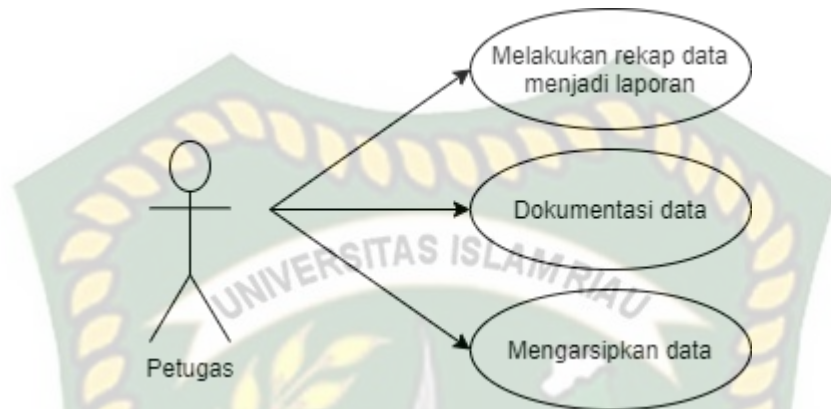
2. Studi Pustaka

Penulis mencari referensi-referensi ke pustaka sebagai acuan penelitian untuk mendukung dalam kasus laporan penelitian. Sumber yang digunakan baik berupa buku, jurnal dan berasal dari peneliti sebelumnya yang berhubungan dengan penelitian.

3.2 Analisis Sistem yang Sedang Berjalan

Sebelum sistem pengelompokan data tunggakan rekening listrik pascabayar dirancang, pihak lapangan yang terkait hanya menyimpan data yang terdapat dalam penelitian ini sebagai dokumentasi laporan. Pihak lapangan melakukan sosialisasi kepada setiap pelanggan pascabayar yang menunggak sehingga banyak

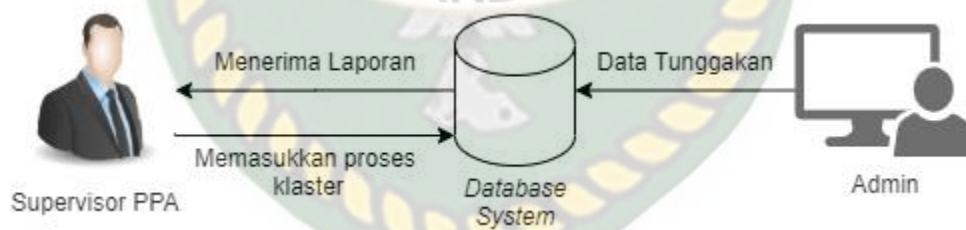
mengeluarkan biaya operasional. Analisa sistem yang sedang berjalan bisa dilihat pada gambar 3.1.



Gambar 3. 1 Analisis Sistem yang Berjalan

3.3 Pengembangan Sistem

Dalam penelitian ini akan dirancang sebuah sistem yang akan memudahkan dalam proses pengelompokan data tunggakan rekening listrik pascabayar. Analisa sistem yang akan diusulkan dapat dilihat pada gambar 3.2.



Gambar 3. 2 Pengembangan Sistem

Berikut adalah gambaran umum sistem yang diusulkan yaitu *admin* melakukan *login* ke sistem dan diarahkan ke menu utama. Menu yang disediakan oleh sistem ini adalah menu penginputan data *training*, proses *cluster*, penginputan data *testing*.

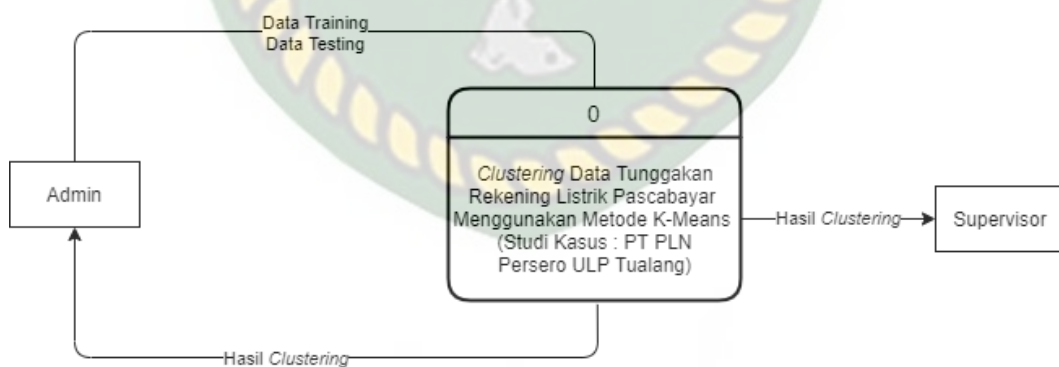
Pada menu *training*, *admin* menginputkan seluruh data pelanggan yang menunggak yang akan dijadikan atribut dalam perhitungan pengelompokan data.

Setelah data masuk kedalam *database* maka sistem dapat melakukan proses pengelompokan terhadap pelanggan. Pengelompokan dilakukan oleh Supervisor PPA agar dapat menghasilkan data pengelompokan pelanggan berdasarkan kriteria yang dimasukkan dalam sistem.

Selanjutnya, hasil dari training akan digunakan sebagai bahan evaluasi oleh Supervisor PPA dan digunakan sebagai pola acuan untuk menentukan kluster pelanggan baru yang menunggak pada tahap *testing*.

3.4.1 Context Diagram

Konteks diagram (*Context Diagram*) digunakan untuk menggambarkan hubungan *input* dan *output* antara sistem dengan entitas luar, suatu diagram konteks selalu memiliki satu poses yang mewakili seluruh sistem. Sistem ini memiliki satu buah eksternal entitas yaitu admin. Konteks diagram dapat dilihat pada gambar 3.3.

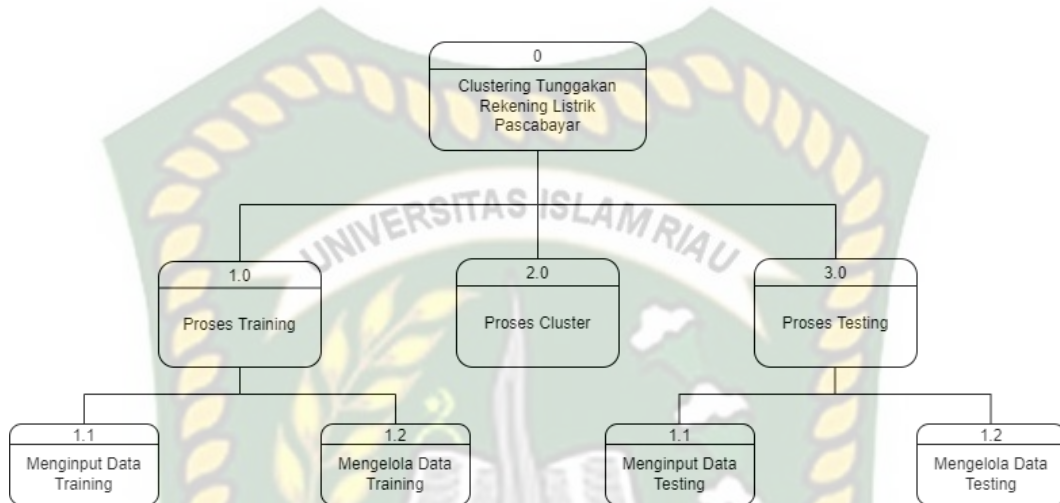


Gambar 3.3 Context Diagram

3.4.2 Hierarchy Chart

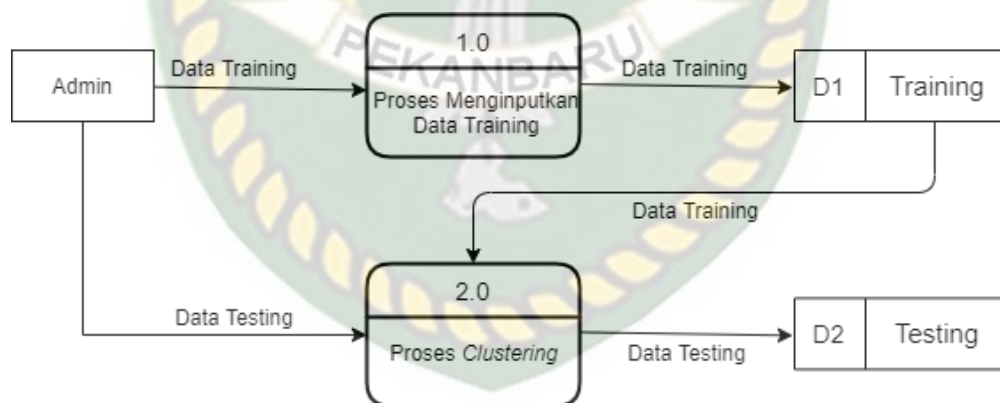
Hierarchy chart merupakan suatu diagram subsistem secara jenjang atau hirarki yang menggambarkan proses yang terdapat dalam sistem dan diuraikan yang

saling berhubungan dengan elemen lainnya dan yang membedakan adalah pada level prosesnya. *Hierarchy chart* sistem yang akan dibangun bisa dilihat pada gambar 3.4.



Gambar 3.4 *Hierarchy Chart*

3.4.3 Data Flow Diagram (DFD) Level 0



Gambar 3.5 DFD Level 0

Bisa dilihat pada gambar 3.5 proses pengolahan data master bertugas mengelola data tunggakan. Data tunggakan diinputkan oleh *admin* kemudian disimpan pada *data store* sebagai data *training*. Selanjutnya *admin* dapat melakukan proses *clustering* dengan mengambil data dari data *training*. *Admin* juga

dapat menginputkan data tunggakan yang akan disimpan pada *data store* sebagai data *testing*. Proses *testing* dilakukan dengan menggunakan acuan dari proses *clustering*.

3.4 Rancangan Desain

3.5.1 Rancangan Desain *Output*

Desain *output* merupakan bentuk sebuah hasil dari proses sistem dalam bentuk laporan. Beberapa desain *output* yang dirancang dalam sistem yang dibangun :

1. *Output* data *training*

Data *training* yang sudah diinputkan seperti pada gambar 3.6.

DATA TRAINING							
No.	ID Pelanggan	Nama Pelanggan	Desa	Daya	Tagihan	kWh	Jam Nyala
9(3)	X(15)	X(50)	X(25)	9(5)	(9(7)	9(5)	9(5)

Gambar 3. 6 *Output* Data *Training*

2. *Output* data *testing*

Data *testing* untuk menginputkan data *testing* seperti pada gambar 3.7.

DATA TESTING							
No.	ID Pelanggan	Nama Pelanggan	Desa	Daya	Tagihan	kWh	Jam Nyala
9(3)	X(15)	X(50)	X(25)	9(5)	(9(7)	9(5)	9(5)

Gambar 3. 7 *Output* Data *Testing*

3. *Output* hasil perhitungan iterasi K-Means

Desain untuk melihat hasil perhitungan K-Means dapat dilihat pada gambar 3.8.

Euclidean Distance Iterasi Ke 9(2)							
No.	ID Pelanggan	Nama Pelanggan	C1	C2	C3	MIN	Cluster
9(3)	X(15)	X(50)	9(7)	9(7)	9(7)	9(7)	X(2)

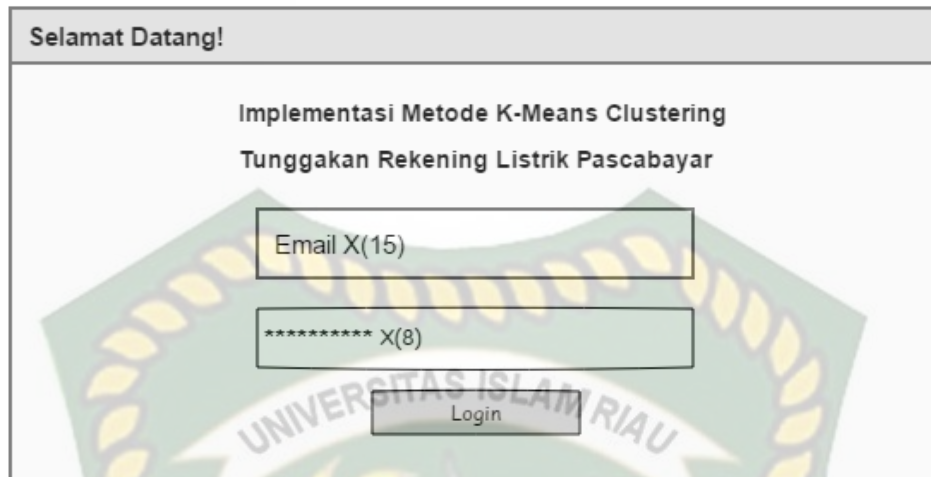
Gambar 3. 8 Output Hasil Iterasi

3.5.2 Rancangan Desain *Input*

Desain *input* adalah bentuk masukan pada sebuah sistem yang akan diproses untuk menghasilkan sebuah informasi.

1. Desain *Input Login*

Desain *input login* merupakan rancangan yang dibuat untuk *user* melakukan login terhadap sistem. Rancangan desain login dapat dilihat pada gambar 3.9.



Selamat Datang!

Implementasi Metode K-Means Clustering
Tunggalan Rekening Listrik Pascabayar

Email X(15)

***** X(8)

Login

Gambar 3.9 *Desain Input Login*

2. *Desain Input Data Training Manual*

Input data training manual ini digunakan untuk data yang ingin diinputkan hanya sedikit. Dengan melakukan *input* data satu persatu sehingga *user* tidak perlu meng-*import* data dari excel. Berikut desain *input data training manual* dapat dilihat pada gambar 3.10.

INPUT DATA TRAINING	
ID Pelanggan	<input type="text" value="X(15)"/>
Nama Pelanggan	<input type="text" value="X(50)"/>
Desa	<input type="text" value="Pilih Desa"/>
Daya	<input type="text" value="9(5)"/>
Tagihan	<input type="text" value="9(7)"/>
kWh	<input type="text" value="9(5)"/>
Jam Nyala	<input type="text" value="9(5)"/>
<input type="button" value="BATAL"/> <input type="button" value="SUBMIT"/>	

Gambar 3. 10 Desain *Input Training* Manual

3. Desain *Input Data Testing* Manual

Input data training manual ini digunakan untuk data yang ingin diinputkan hanya sedikit. Dengan melakukan *input* data satu persatu sehingga *user* tidak perlu meng-*import* data dari excel. Berikut desain *input data training* manual dapat dilihat pada gambar 3.11.

INPUT DATA TESTING

ID Pelanggan	<input type="text" value="X(15)"/>
Nama Pelanggan	<input type="text" value="X(50)"/>
Desa	<input type="text" value="Pilih Desa"/>
Daya	<input type="text" value="9(5)"/>
Tagihan	<input type="text" value="9(7)"/>
kWh	<input type="text" value="9(5)"/>
Jam Nyala	<input type="text" value="9(5)"/>
<input type="button" value="BATAL"/> <input type="button" value="SUBMIT"/>	

Gambar 3. 11 Desain *Input Testing* Manual

3.5.3 Perancangan *Database*

3.5.3.1 Desain *Database*

1. Tabel *User*

Nama *Database* : *tunggakan_listrik*

Nama Tabel : *User*

Tabel 3. 1 Tabel *User*

No	Field	Data Type	Size	Keterangan
1	Email	Varchar	20	Primary Key
2	Password	Varchar	10	

2. Tabel Data *Training*

Nama *Database* : tunggakan_listrik

Nama Tabel : data_bisnis

Tabel 3. 2 Tabel *Training*

No	Field	Data Type	Size	Keterangan
1	Id_bisnis	Int	-	Primary Key
2	Idpel	Char	15	
3	nama	Varchar	50	
4	desa	Varchar	50	
5	daya	Int	-	
6	rptag	Int	-	
7	pemkwh	Int	-	
8	jamnyala	Int	-	

3. Tabel Data *Testing*

Nama *Database* : tunggakan_listrik

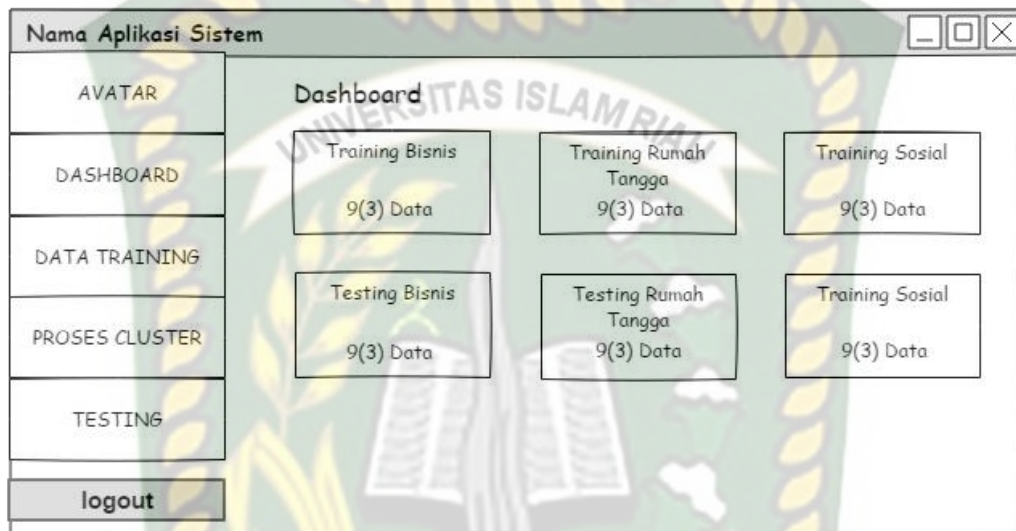
Nama Tabel : datatest

Tabel 3. 3 Tabel Data *Testing*

No	Field	Data Type	Size	Keterangan
1	Id_dataset	Int	-	Primary Key
2	Idpel	Char	15	
3	nama	Varchar	50	
4	desa	Varchar	50	
5	daya	Int	-	
6	rptag	Int	-	
7	pemkwh	Int	-	
8	jamnyala	Int	-	
9	label	enum	'C1', 'C2', 'C3'	

3.5.4 Desain Antarmuka

Desain antarmuka merupakan bagian dari sistem yang akan digunakan sebagai media interaksi antara sistem dengan pengguna. Adapun desain antarmuka dari sistem terdapat pada gambar 3.12.



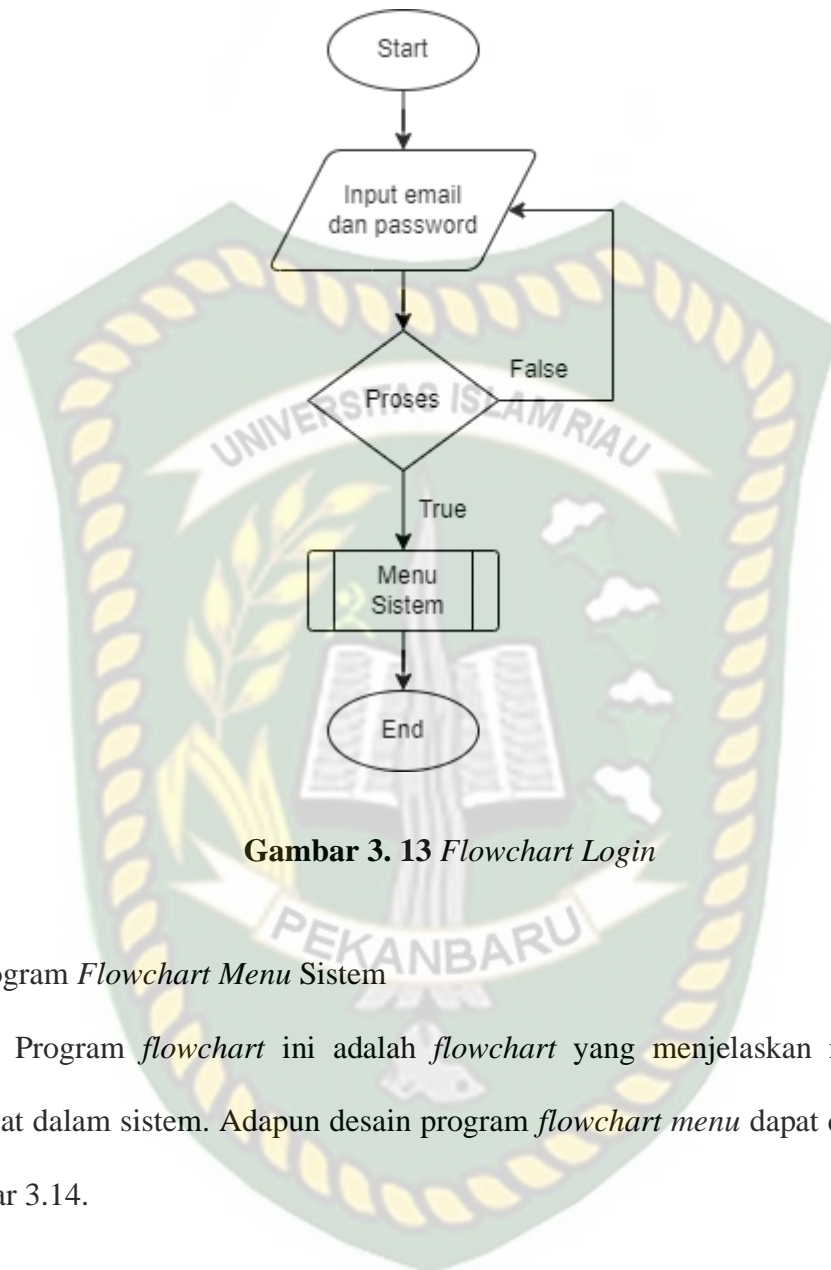
Gambar 3. 12 Desain Antarmuka

3.5.5 Rancangan Logika Program

Desain logika program menggambarkan logika program yang akan dibuat menunjukkan alat bantu *flowchart* yang bertujuan untuk menunjukkan alur sistem maupun program, mulai dari data *input* sampai hasil *output*.

1. Program *Flowchart Login*

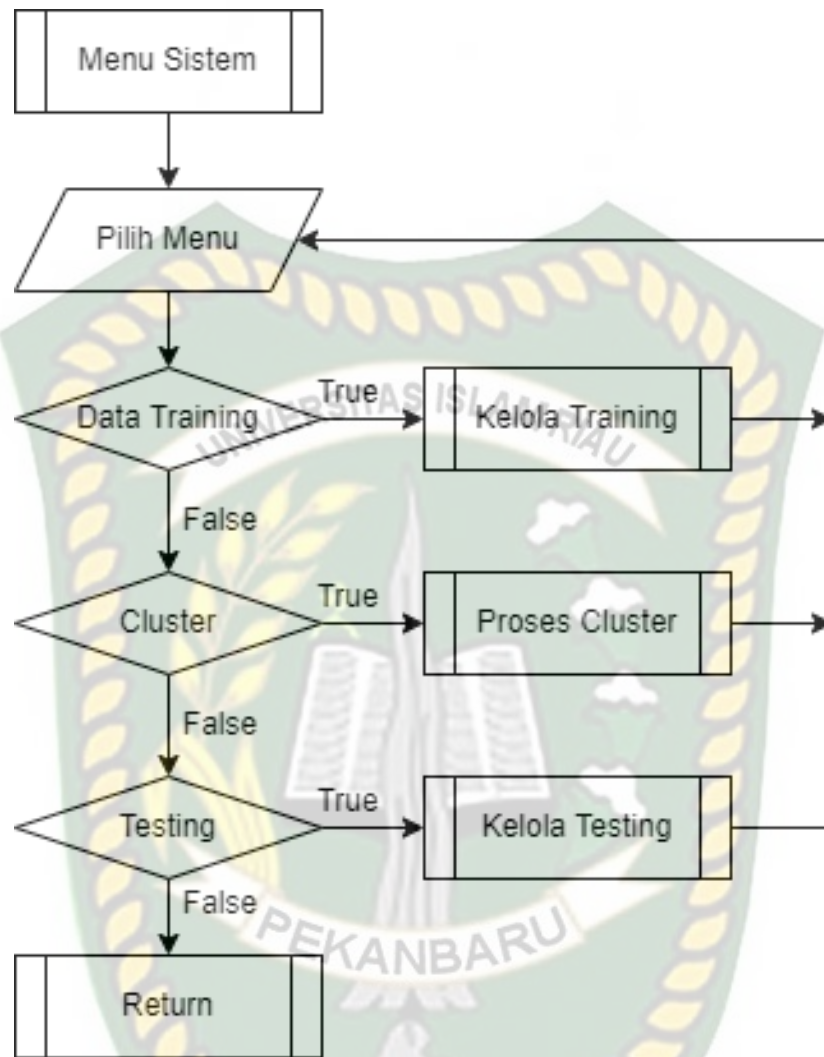
Program *flowchart login* adalah *flowchart* yang menjelaskan bagaimana proses pengguna untuk masuk kedalam sistem. Adapun desain program *flowchart login* dapat dilihat pada gambar 3.13.



Gambar 3. 13 *Flowchart Login*

2. Program *Flowchart Menu Sistem*

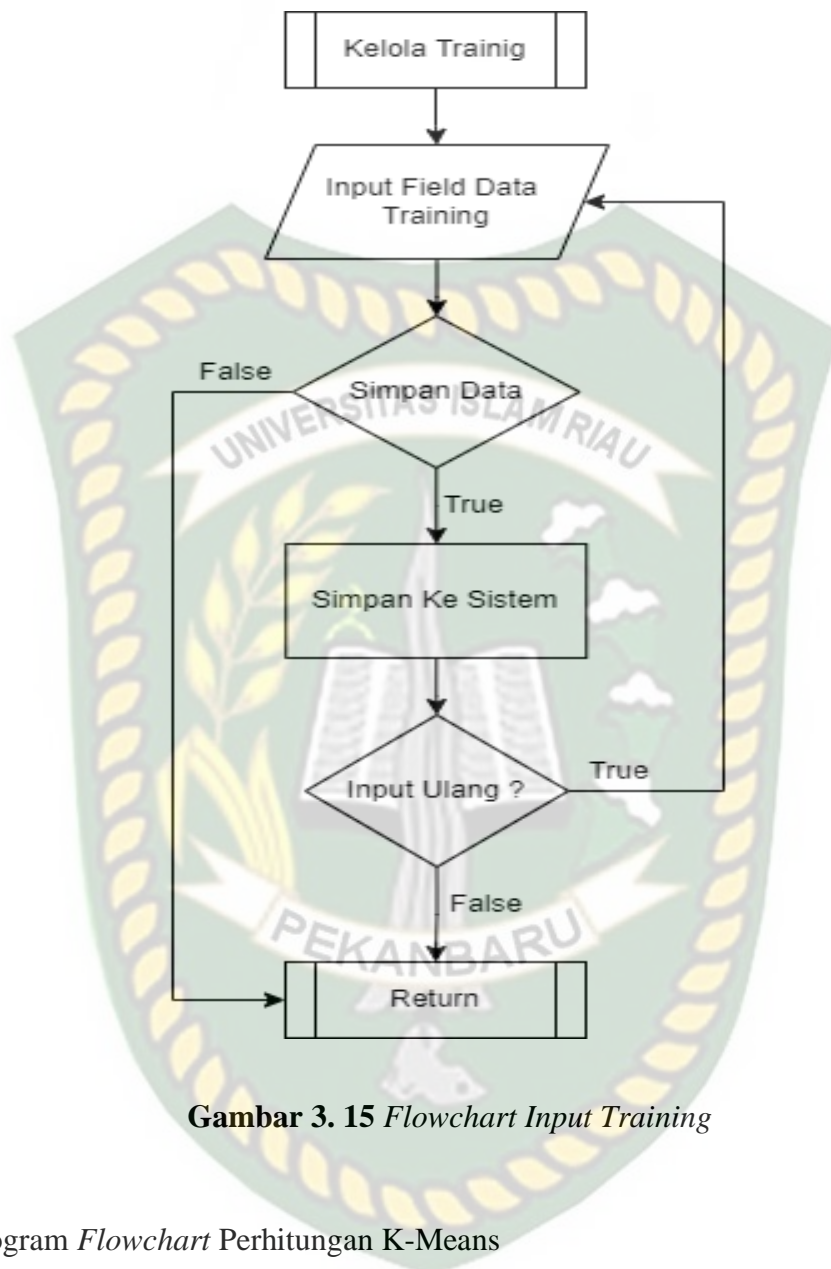
Program *flowchart* ini adalah *flowchart* yang menjelaskan menu yang terdapat dalam sistem. Adapun desain program *flowchart menu* dapat dilihat pada gambar 3.14.



Gambar 3. 14 *Flowchart Menu Sistem*

3. Program *Flowchart Input Data Training*

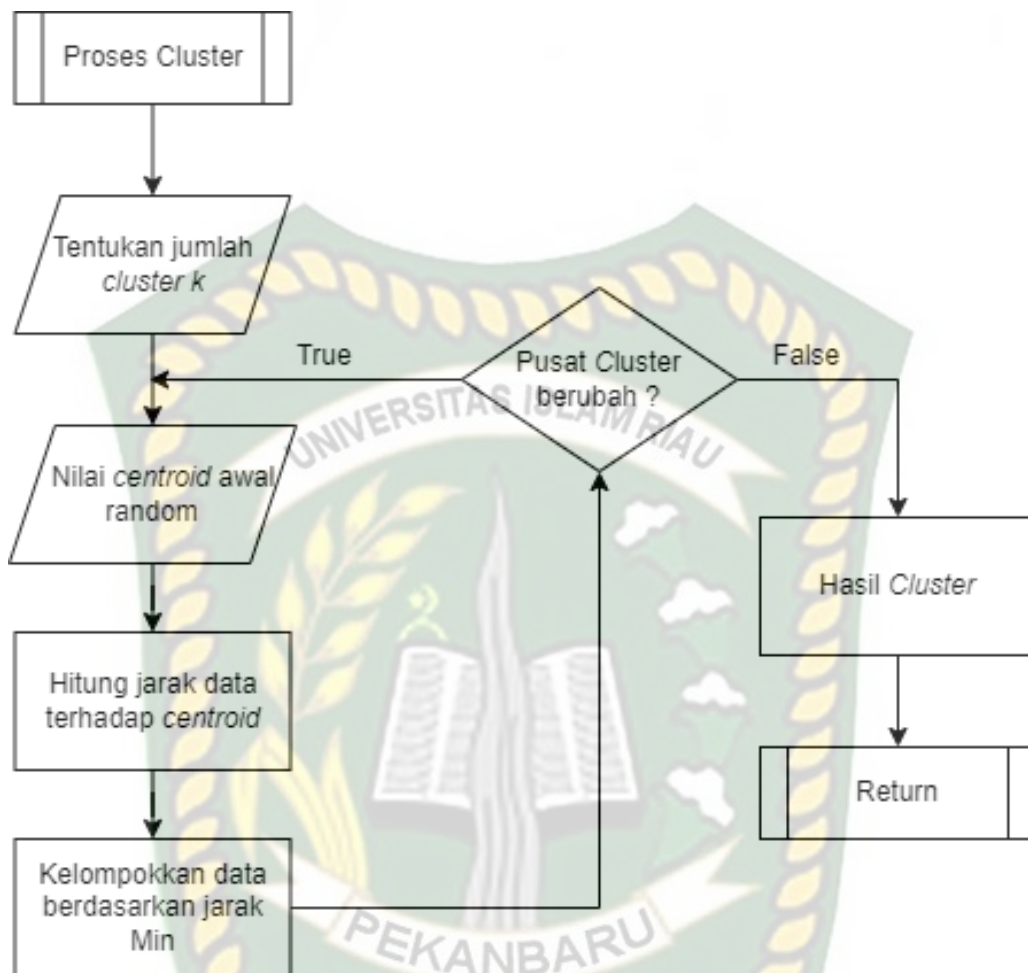
Program *flowchart Input data training* adalah *flowchart* yang menjelaskan bagaimana proses *input data training* yang akan tersimpan kedalam sistem. Adapun desain program *flowchart* ini dapat dilihat pada gambar 3.15.



Gambar 3. 15 *Flowchart Input Training*

4. Program *Flowchart* Perhitungan K-Means

Program *flowchart* ini adalah *flowchart* yang menjelaskan bagaimana proses perhitungan K-Means dan menghasilkan *output cluster*. Adapun desain program *flowchart* ini dapat dilihat pada gambar 3.16.



Gambar 3. 16 *Flowchart* Algoritma K-Means

BAB IV

HASIL DAN PEMBAHASAN

4.1 Pengujian *Black Box*

Pengujian *black box* (*black box testing*) adalah salah satu metode pengujian perangkat lunak yang berfokus pada sisi fungsionalitas, khususnya pada input dan output pada aplikasi untuk menentukan apakah aplikasi tersebut sudah sesuai dengan yang di harapkan.

4.1.1 Pengujian *Form Training*

Untuk dapat melakukan pengolahan data *training* pada sistem, pengguna harus masuk ke sistem dan memilih menu data *training*. Pengguna memasukkan *id pelanggan*, *nama* dan *data lainnya* yang telah terdaftar ke sistem. Berikut tampilan halaman *form training* sistem ini :

The image shows a web application interface. A modal form titled "Tambah Data Bisnis" is open over a "DATA TRAINING" table. The form has the following fields:

- ID* (text input): ID Pelanggan
- Nama* (text input): ex. Balqis
- Desa* (dropdown menu): -Pilih Desa-
- Daya* (text input): ex. 900
- Tagihan* (text input): Rp.
- kWh* (text input): ex. 300
- Jam Nyala* (text input): ex. 125

At the bottom of the form are "Batal" and "Submit" buttons. The background table has columns: No., ID, Pelanggan, Pemakaian (kWh), Jam Nyala, and Aksi. The footer of the application reads "Copyright © 2021 Balqis Malpasha, Teknik Informatika."

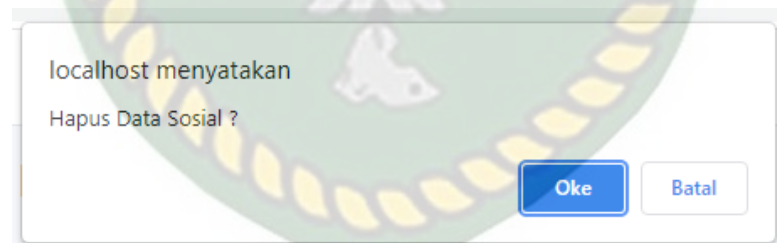
Gambar 4. 1 Pengujian *Form Training*

Pada gambar 4.1 dijelaskan bahwa setelah memasukkan data, maka data akan tersimpan pada *database*. Dapat dilihat pada gambar 4.2.

No.	ID	Petanggan	Nama	Desa	Days	Rptag	Pemakaian (kWh)	Jam Nyala	Aksi
1	181500209607	PASAR TUAH SERUMPUN	PASAR TUAH SERUMPUN	PERAWANG	3300	Rp307,310.00	251	73	[Edit] [Hapus]
2	181500000929	SYAHRIL EFENDI	PERAWANG	PERAWANG	900	Rp114,675.00	209	230	[Edit] [Hapus]
3	181500001002	DESMAN	PERAWANG BARAT	PERAWANG BARAT	900	Rp105,900.00	192	213	[Edit] [Hapus]
4	181500000720	MUSNIMAR II	PERAWANG BARAT	PERAWANG BARAT	5500	Rp1,182,170.00	977	178	[Edit] [Hapus]
5	181500018115	DARVUL ISMAN	MARE DAN	MARE DAN	1300	Rp47,133.00	82	63	[Edit] [Hapus]

Gambar 4. 2 Tampilan Data *Training*

Pada gambar 4.2 menyatakan bahwa hasil penginputan dapat di ubah atau di hapus. Jika data akan dihapus, klik icon hapus dan akan muncul pesan seperti berikut ini :



Gambar 4. 3 Tampilan Pesan Hapus

Tabel 4. 1 Kesimpulan Pengujian *Form Training*

No.	Komponen yang Diuji	Skenario Pengujian	Hasil yang Diharapkan	Hasil
1	<i>Form Training</i>	Mengosongkan pada semua <i>field</i> , klik submit.	Sistem menolak memulai pemrosesan yang diinput.	[✓] Sesuai Harapan [] Tidak Sesuai Harapan
		Menginputkan nama dan mengosongkan <i>field</i> training lainnya.	Sistem menolak	[✓] Sesuai Harapan [] Tidak Sesuai Harapan
		Mengisi pada semua <i>field</i> , klik submit.	Sistem menyimpan data yang diinput.	[✓] Sesuai Harapan [] Tidak Sesuai Harapan
		Mengisi semua <i>field</i> , klik batal.	Sistem akan menutup <i>form training</i> .	[✓] Sesuai Harapan [] Tidak Sesuai Harapan
2	Mengklik icon “Edit”	Mengedit data <i>training</i>	Sistem akan kembali pada <i>form training</i>	[✓] Sesuai Harapan [] Tidak Sesuai Harapan
3	Mengklik icon “hapus”	Menghapus data <i>training</i>	Sistem akan menghapus dan kembali pada data <i>training</i> .	[✓] Sesuai Harapan [] Tidak Sesuai Harapan

4.1.2 Pengujian Proses K-Means

Pengujian selanjutnya yaitu *form* proses K-Means yang mana dapat dilihat pada gambar 4.4. Pada *form* proses k-means yang harus ditampilkan yaitu hasil dari proses data *training* yang dimasukkan sebelumnya. *Form* ini harus diinputkan dengan benar sesuai dengan formatnya masing-masing.

CLUSTERING Home / Cluster / Bisnis

Centroid Awal

No.	ID Pelanggan	Nama	Desa	Daya	Rptag	Pemakaian (kWh)	Jam Nyala
1	18150000720	YUSNIMAR.II	PERAWANG BARAT	5500	Rp1,182,170.00	977	178
2	18150003510	SYAFRIZAL	PINANG SCDATANG BARAT	2200	Rp315,850.00	261	119
3	181500112386	SUTARMAN	PERAWANG	2200	Rp618,310.00	511	232

Gambar 4. 4 Pusat Cluster Awal

Euclidean Distance Iterasi Ke-8

Excel PDF Print Search:

No. ↑↓	ID	Nama	C1	C2	C3	Minimum	Cluster
1	181500205607	PASAR TUAH SERUMPUN	Rp1,320,060.41	Rp140,980.50	Rp101,723.41	Rp101,723.41	c3
2	18150000929	SYAHRIAL EFFENDI I	Rp1,512,736.89	Rp51,702.26	Rp294,379.77	Rp51,702.26	c2
3	181500001602	DESMAN	Rp1,521,431.74	Rp66,396.44	Rp303,074.72	Rp66,396.44	c2
4	18150000720	YUSNIMAR.II	Rp145,237.48	Rp1,015,805.79	Rp773,125.47	Rp145,237.48	c1
5	181500019115	DARYUL ISMAN	Rp1,540,276.13	Rp79,239.65	Rp221,620.56	Rp79,239.65	c2
6	181500021443	SALIM	Rp1,408,525.43	Rp52,525.19	Rp193,159.04	Rp52,525.19	c2
7	181500001610	ISLAMU	Rp1,591,811.35	Rp130,775.45	Rp373,455.50	Rp130,775.45	c2

Gambar 4. 5 Hasil Clustering Tunggalan

Tabel 4. 2 Kesimpulan Pengujian *Form* Proses *K-Means*

No.	Komponen yang Diuji	Skenario Pengujian	Hasil yang Diharapkan	Hasil
1	Mengklik tombol “Cluster Bisnis”	Menghitung proses k-means	Sistem akan menampilkan hasil <i>cluster</i> proses k-means	[<input checked="" type="checkbox"/>] Sesuai Harapan [<input type="checkbox"/>] Tidak Sesuai Harapan
2	Mengklik tombol “Cluster Rumah Tangga”	Menghitung proses k-means	Sistem akan menampilkan hasil <i>cluster</i> proses k-means	[<input checked="" type="checkbox"/>] Sesuai Harapan [<input type="checkbox"/>] Tidak Sesuai Harapan
3	Mengklik tombol “Cluster Sosial”	Menghitung proses k-means	Sistem akan menampilkan hasil <i>cluster</i> proses k-means	[<input checked="" type="checkbox"/>] Sesuai Harapan [<input type="checkbox"/>] Tidak Sesuai Harapan

4.1.3 Pengujian Proses Perhitungan Metode K-Means

Pengujian selanjutnya yaitu proses menampilkan data perhitungan k-means dari sistem yang dibangun. Hasil proses perhitungan akan menampilkan data dari iterasi-iterasi yang dihasilkan pada proses perhitungan jumlah iterasi yang dilakukan tergantung data yang diproses. Berikut ini adalah gambar hasil proses perhitungan metode K-means di program.

Euclidean Distance Iterasi Ke-8

Excel PDF Print Search:

No.	ID	Nama	C1	C2	C3	Minimum	Manual	Cluster
1	181500153682	WAKIDI	Rp3,512,719.22	Rp3,500,045.40	Rp3,651,048.00	Rp3,500,045.40	C2	c2
2	181500055329	TKHIRUL	Rp3,500,127.53	Rp3,513,753.00	Rp3,757,944.75	Rp3,500,127.53	C1	c1
3	181500055585	ZUBIR	Rp3,669,312.75	Rp3,556,428.38	Rp3,844,218.19	Rp3,666,428.38	C2	c2
4	181500057252	SAMULFI PANJAITAN	Rp2,749,237.62	Rp2,710,436.17	Rp2,821,772.12	Rp2,710,436.17	C3	c2
5	181500062433	ERWIN SIREGAR	Rp2,502,683.71	Rp2,505,397.15	Rp2,782,703.83	Rp2,502,683.71	C1	c1
6	181500055386	DAMBANG	Rp6,602,643.03	Rp6,539,536.58	Rp6,406,385.06	Rp6,406,385.06	C3	c3
7	181500138678	YUJIN ARMINDRA	Rp2,600,981.88	Rp2,523,707.39	Rp2,957,521.83	Rp2,600,981.88	C1	c1
8	181500016062	RAZAK	Rp3,205,389.96	Rp3,201,391.87	Rp3,401,090.92	Rp3,201,391.87	C2	c2

Copyright © 2021 Balqis Maipasha, Teknik Informatika.

Gambar 4. 6 Pengujian Hasil Perhitungan Proses *K-means*

4.1.4 Kesimpulan Pengujian *Black Box*

Dengan dilakukannya proses pengujian *black box* ini dapat disimpulkan bahwa setiap data yang telah diinputkan kedalam sistem harus sesuai dengan format yang benar sesuai dengan sistem yang telah dibuat. Kolom yang diberikan tanda bintang merah tidak boleh dikosongkan. Apabila ada kesalahan dalam penginputan data kedalam sistem maka sistem akan menolak dan tidak dapat disimpan kedalam sistem. Jika *form* diisi dengan benar dan sesuai maka data dapat disimpan ke sistem dan sistem akan berjalan sesuai yang diharapkan.

4.2 Pengujian *Whitebox*

Pengujian *whitebox* merupakan proses pengujian berdasarkan perhitungan manual yang sudah dilakukan dan akan disamakan dengan proses perhitungan dari sistem untuk menyesuaikan hasil dari sistem dengan manual.

4.2.1 Perhitungan Manual Metode K-Means

Perhitungan metode K-means dilakukan dengan data dari bulan April tahun 2021. Tahap perhitungan metode ini dimulai dari penentuan pusat *cluster* secara random untuk perhitungan jarak pada iterasi pertama dari data yang dikumpulkan. Proses iterasi akan berhenti apabila pusat *cluster* iterasi sebelumnya dan sesudahnya mencapai nilai yang konvergen. Berikut ini data tunggakan pada bulan April 2021.

4.2.1.1 Tarif Bisnis

Tabel 4. 3 Data *Training* Tarif Bisnis

No.	Nama	Desa	Daya	Rptag	kWh	Jam Nyala
1	PASAR TUAH SERUMPUN	PERAWANG	3500	Rp307,340.00	254	73
2	SYAHRIAL EFENDI. I	PERAWANG	900	Rp114,675.00	209	232
3	DESMAN	PERAWANG BARAT	900	Rp105,980.00	192	213
4	YUSNIMAR.II	PERAWANG BARAT	5500	Rp1,182,170.00	977	178
5	DARYULISMAN	MAREDAN	1300	Rp87,133.00	82	63
6	SAHAR	PERAWANG BARAT	1300	Rp218,896.00	206	158
7	ISLAMI	PERAWANG	1300	Rp35,597.00	15	12
8	NASRUL	PERAWANG	2200	Rp72,600.00	32	15
9	SYAFRIZAL.	PINANG SEBATANG BARAT	2200	Rp315,810.00	261	119
10	DON TETELES BIE LEE	PERAWANG	1300	Rp279,464.00	263	202
11	ELIZA ALIAS GIOK TJU 1	PERAWANG	1300	Rp362,347.00	341	262
12	JON GUNARDI.	PERAWANG BARAT	1300	Rp85,008.00	80	62
13	ADE ROSMADINA	PERAWANG BARAT	5500	Rp2,680,150.00	2215	403
14	SUDIRMAN SITORUS	PERAWANG	1300	Rp463,782.00	442	340
15	KHO.KA.SING	MAREDAN	4400	Rp539,660.00	446	101
16	M.BASIR TALIB.R.HUSD	PERAWANG	4400	Rp352,110.00	291	66
17

88	RIKO ADRIANSYAH	PINANG SEBATANG BARAT	1300	Rp426,103.00	401	308

Iterasi Ke-1 :

- a. Menentukan pusat awal *Cluster*

Nilai Random awal :

Tabel 4. 4 Titik Pusat *Cluster* Awal

C	DAYA	RPTAG	PEMKWH	JAM NYALA
C1	5500	1182170	977	178
C2	2200	315810	261	119
C3	2200	618310	511	232

- b. Perhitungan jarak pusat *cluster*

Untuk mengukur jarak antara data dengan pusat *cluster* digunakan *Euclidean distance* dengan rumus sebagai berikut :

Berikut contoh perhitungan jarak dari data ke-1 terhadap pusat *cluster* :

$$d(x_1, c_1) = \sqrt{(5500 - 3500)^2 + (1182170 - 307340)^2 + (977 - 254)^2 + (178 - 73)^2} = 874833$$

$$d(x_1, c_2) = \sqrt{(2200 - 3500)^2 + (315810 - 307340)^2 + (261 - 254)^2 + (119 - 73)^2} = 8569$$

$$d(x_1, c_3) = \sqrt{(2200 - 3500)^2 + (618310 - 307340)^2 + (511 - 254)^2 + (232 - 73)^2} = 310973$$

Dan seterusnya dilanjutkan untuk data ke 2, 3,..n. Kemudian akan didapatkan matrik jarak untuk iterasi ke-1.

Tabel 4. 5 Perhitungan Jarak Iterasi-1

No.	Nama	C1	C2	C3	Cluster
1	PASAR TUAH SERUMPUN	Rp874.833	Rp8.569	Rp310.973	C2
2	SYAHRIAL EFENDI. I	Rp1.067.505	Rp201.139	Rp503.637	C2
3	DESMAN	Rp1.076.200	Rp209.834	Rp512.332	C2
4	YUSNIMAR.II	Rp0	Rp866.367	Rp563.870	C1
5	DARYULISMAN	Rp1.095.045	Rp228.679	Rp531.178	C2
6	SAHAR	Rp963.283	Rp96.918	Rp399.415	C2
7	ISLAMI	Rp1.146.581	Rp280.215	Rp582.714	C2
8	NASRUL	Rp1.109.575	Rp243.210	Rp545.710	C2
9	SYAFRIZAL.	Rp866.367	Rp0	Rp302.500	C2
10	DON TETELES BIE LEE	Rp902.716	Rp36.357	Rp338.847	C2
11	ELIZA ALIAS GIOK TJU 1	Rp819.834	Rp46.546	Rp255.965	C2
12	JON GUNARDI.	Rp1.097.170	Rp230.804	Rp533.303	C2
13	ADE ROSMADINA	Rp1.497.981	Rp2.364.343	Rp2.061.843	C1
14	SUDIRMAN SITORUS	Rp718.400	Rp147.975	Rp154.531	C2
15	KHO.KA.SING	Rp642.511	Rp223.861	Rp78.681	C3
16	M.BASIR TALIB.R.HUSD	Rp830.061	Rp36.367	Rp266.209	C2
17

88	RIKO ADRIANSYAH	Rp756.079	Rp110.297	Rp192.209	C2

Setiap kolom pada matrik menunjukkan nilai jarak data terhadap pusat *cluster*. Baris pertama pada matrik menunjukkan nilai jarak data terhadap titik pusat *cluster* pertama, dan baris kedua pada matriks menunjukkan nilai jarak data terhadap titik pusat *cluster* kedua begitu juga dengan baris ketiga.

c. Pengelompokan data

Jarak dari hasil perhitungan yang sudah dilakukan, akan dibandingkan dan dipilih jarak terdekat atau nilai terkecil dari ketiga hasil perhitungan jarak setiap *cluster*, jarak ini menunjukkan bahwa data tersebut berada dalam satu kelompok dengan *cluster* terdekat.

Tabel 4. 6 Pengelompokan *Cluster* Iterasi-1

No.	C1	C2	C3
1		*	
2		*	
3		*	
4	*		
5		*	
6		*	
7		*	
8		*	
9		*	
10		*	
11		*	
12		*	
13	*		
14		*	
15			*
16		*	
17	.	.	.
	.	.	.
	.	.	.
88		*	

d. Perhitungan titik *centroid* baru

Selanjutnya menghitung centroid atau pusat *cluster* baru setiap anggota klasternya dengan menjumlahkan tiap atribut yang termasuk dalam satu klaster dan mencari rata-rata hasil dari tiap-tiap atribut pada masing-masing cluster.

Tabel 4. 7 Pusat *Cluster* Baru

	DAYA	TAGIHAN	PEMKWH	JAMNYALA
C1	9800	2906053	1966	287
C2	1663	221730	233	190
C3	2044	546014	478	275

Iterasi ke-2 :

Ulangi langkah kedua (b) hingga data tidak mengalami perubahan. Berikut hasil perhitungan jarak iterasi kedua :

Tabel 4. 8 Perhitungan Jarak Iterasi-2

No.	Nama	C1	C2	C3	Cluster
1	PASAR TUAH SERUMPUN	Rp2.598.721	Rp85.630	Rp238.679	C2
2	SYAHRIAL EFENDI. I	Rp2.791.392	Rp107.058	Rp431.341	C2
3	DESMAN	Rp2.800.087	Rp115.752	Rp440.036	C2
4	YUSNIMAR.II	Rp1.723.888	Rp960.448	Rp636.165	C3
5	DARYULISMAN	Rp2.818.933	Rp134.598	Rp458.882	C2
6	SAHAR	Rp2.687.171	Rp2.857	Rp327.119	C2
7	ISLAMI	Rp2.870.469	Rp186.134	Rp510.418	C2
8	NASRUL	Rp2.833.464	Rp149.131	Rp473.414	C2
9	SYAFRIZAL.	Rp2.590.254	Rp94.082	Rp230.204	C2
10	DON TETELES BIE LEE	Rp2.626.603	Rp57.735	Rp266.551	C2
11	ELIZA ALIAS GIOK TJU 1	Rp2.543.720	Rp140.618	Rp183.669	C2
12	JON GUNARDI.	Rp2.821.058	Rp136.723	Rp461.007	C2
13	ADE ROSMADINA	Rp225.944	Rp2.458.424	Rp2.134.139	C1
14	SUDIRMAN SITORUS	Rp2.442.286	Rp242.052	Rp82.236	C3
15	KHO.KA.SING	Rp2.366.399	Rp317.942	Rp6.779	C3
16	M.BASIR TALIB.R.HUSD	Rp2.553.949	Rp130.409	Rp193.919	C2
17

88	RIKO ADRIANSYAH	Rp2.479.965	Rp204.373	Rp119.913	C3

Langkah berikutnya, melakukan perbandingan perbandingan hasil *cluster* pada iterasi pertama dengan kedua. Jika nilai anggota *cluster* memiliki nilai yang konvergen/sama maka proses iterasi berhenti. Iterasi dihentikan dan hasil *cluster* tersebut merupakan hasil *cluster* yang sudah valid.

Tabel 4. 9 Perbandingan Iterasi

No.	Iterasi 1	Iterasi 2
1	C2	C2
2	C2	C2
3	C2	C2
4	C1	C3
5	C2	C2
6	C2	C2
7	C2	C2
8	C2	C2
9	C2	C2
10	C2	C2
11	C2	C2
12	C2	C2
13	C1	C1
14	C2	C3
15	C3	C3
16	C2	C2
17	.	.
88	C2	C3

Kesimpulannya, jika nilai anggota *cluster* sebelumnya memiliki nilai yang konvergen dengan nilai anggota *cluster* setelahnya maka proses iterasi dihentikan. Jika nilai anggota *cluster* belum konvergen maka ulangi langkah sebelumnya (d) untuk menghitung pusat *cluster* baru.

No.	ID	Nama	C1	C2	C3	Minimum	Cluster
1	181500205607	PASAR TUAH SERUMPUN	Rp3,908,974.77	Rp39,162.46	Rp1,162,440.05	Rp39,162.46	c2
2	18150000929	SYAHRIAL EFENDI I	Rp5,101,644.69	Rp153,545.67	Rp1,655,114.35	Rp153,545.67	c2
3	18150001602	DESMAN	Rp6,110,339.68	Rp162,240.56	Rp1,663,809.25	Rp162,240.56	c2
4	18150000720	YUSNIMAR,II	Rp3,034,142.29	Rp913,959.62	Rp587,609.55	Rp587,609.55	c3
5	181500019115	DARYULISMAN	Rp5,129,185.82	Rp181,086.12	Rp1,682,654.29	Rp181,086.12	c2
6	181500021443	SAHAR	Rp5,997,423.08	Rp49,324.30	Rp1,550,892.65	Rp49,324.30	c2
7	18150001610	ISLAMI	Rp5,180,721.73	Rp232,622.08	Rp1,734,189.82	Rp232,622.08	c2
8	18150000227	NASRUL	Rp5,143,716.91	Rp195,619.33	Rp1,697,183.41	Rp195,619.33	c2
9	181500073113	SYAFRIZAL	Rp4,505,507.35	Rp47,594.05	Rp1,453,975.48	Rp47,594.05	c2
10	18150001707	DON TETELES BIE LEE	Rp5,936,855.20	Rp11,253.12	Rp1,490,325.38	Rp11,253.12	c2

Gambar 4. 7 Hasil Perhitungan Sistem

Dari hasil *training* diatas dari 88 pelanggan terdapat 3 pelanggan di *cluster* 1, 78 pelanggan di *cluster* 2 dan 7 pelanggan di *cluster* 3. Hasil *cluster* tersebut akan diuji performanya dengan menggunakan RapidMiner dengan metode perhitungan Davies Bouldin yang menghasilkan 0,117 yang artinya 11,7% kemiripan data setiap *cluster*.

Selanjutnya, pengujian hasil dari sistem akan dibandingkan dengan hasil manual dengan menggunakan 20 data testing pada tarif bisnis. Berikut hasil perbandingan pengujian dari sistem dengan manual.

No.	ID	Nama	C1	C2	C3	Minimum	Manual	Cluster
1	18150004795	MUNIR	Rp5,193,217.19	Rp15,591.60	Rp1,252,085.01	Rp15,591.60	C2	c2
2	181500153371	AZRI	Rp5,417,424.36	Rp208,621.91	Rp1,476,291.74	Rp208,621.91	C2	c2
3	181500230390	SUPARTO 02	Rp2,795,157.78	Rp2,413,446.40	Rp1,145,776.43	Rp1,145,776.43	C3	c3
4	181500097992	HERMANSYAH	Rp5,151,725.35	Rp57,030.01	Rp1,210,643.31	Rp57,030.01	C2	c2
5	181500145974	NIRWAN 2	Rp4,950,955.55	Rp257,847.11	Rp1,009,823.29	Rp257,847.11	C2	c2
6	181500187051	BUN SEN	Rp0.00	Rp5,208,802.56	Rp3,941,132.80	Rp0.00	C1	c1
7	18150065917	IRWAN	Rp5,184,484.69	Rp24,322.03	Rp1,243,351.99	Rp24,322.03	C2	c2
8	181500038336	M. RASYID	Rp5,467,623.77	Rp258,624.14	Rp1,526,490.99	Rp258,624.14	C2	c2
9	181500101938	JONI	Rp4,821,486.06	Rp387,316.99	Rp880,354.33	Rp387,316.99	C3	c2
10	181500031596	ZA NURAS	Rp5,133,711.43	Rp75,093.69	Rp1,192,579.44	Rp75,093.69	C2	c2

Gambar 4. 8 Hasil Testing Sistem

Dari 20 data *testing* pada tarif bisnis, hasilnya mendapatkan perbedaan 1 data. Sehingga perbandingannya adalah $19/20 * 100\% = 95\%$ artinya sistem ini dapat digunakan dengan baik.

4.2.1.2 Tarif Sosial

Tabel 4. 10 Data Training Tarif Sosial

No.	Nama	Daya	Rptag	kWh	Jam Nyala
1	TK AR RAHMAN	1300	27258	18	14
2	MESJID NURUL IMAN	1300	87226	112	86
3	SMK YAMATU TUALANG	10600	1437480	1452	137
4	SMPN 39 SIAK	7700	464310	469	61
5	MUSYALLA MUHAMMADIYA	900	12799	23	26
6	MUSHALLA ISTIQOMAH	900	19289	43	48
7	SMAN 1 SUNGAI MANDAU	16500	781110	789	48
8	NIRWAN 1	900	34353	15	17
9	SD NEGERI 04 PERAWANG BAR	1300	421331	541	416
10	SLTP 4 MINAS.	3500	729630	737	211
11	SD 020	900	43021	106	118
12	MUSALLA DARUSADAH	1300	239092	307	236
13	MESJID DARUSSALAM	2200	250800	300	136
14	MULL.DT.MARAJ0	450	56859	165	367
15	MUSH.DARUSALAM	450	51711	152	338

16	SD NO.056	450	71115	201	447
17	MUSH TANAH LAPANG	450	22011	77	171
18	MUSH. AMANAH	450	4910	18	40
19	M.NURUL IKLAS	450	11781	48	107
20	MUSH. M. R A S U L	450	38247	118	262
21	MUSH ISTIQOMAH	2200	61028	58	26
22	MUSALLA NURUL FALAH	450	4369	14	31
23	MESJID DAL FALAH	900	39457	97	108
24	SD. 047	900	18964	42	47
25
60	SMPS LPM DESA MAYANG PONG	1300	56852	73	56

Iterasi Ke-1 :

- a. Menentukan pusat awal *Cluster*

Nilai Random awal :

Tabel 4. 11 Pusat *Cluster* Awal

C	DAYA	RPTAG	PEMKWH	JAM NYALA
C1	900	135685	340	378
C2	16500	781110	789	48
C3	450	4910	18	40

- b. Perhitungan jarak pusat *cluster*

Untuk mengukur jarak antara data dengan pusat *cluster* digunakan

Euclidean distance dengan rumus sebagai berikut :

Berikut contoh perhitungan jarak dari data ke-1 terhadap pusat *cluster* :

$$d(x_1, c_1) = \sqrt{(900 - 1300)^2 + (135685 - 27258)^2 + (340 - 18)^2 + (378 - 14)^2} = 108429$$

$$d(x_1, c_2) = \sqrt{(16500 - 1300)^2 + (781110 - 27258)^2 + (789 - 18)^2 + (48 - 14)^2} = 754006$$

$$d(x_1, c_3) = \sqrt{(450 - 1300)^2 + (4910 - 27258)^2 + (18 - 18)^2 + (40 - 14)^2} = 22364$$

Dan seterusnya dilanjutkan untuk data ke 2, 3,..n. Kemudian akan didapatkan matrik jarak untuk iterasi ke-1.

Tabel 4. 12 Perhitungan Jarak Iterasi-1

No.	Nama	C1	C2	C3	Cluster
1	TK AR RAHMAN	Rp108.429	Rp754.006	Rp22.364	C3
2	MESJID NURUL IMAN	Rp48.462	Rp694.051	Rp82.320	C1
3	SMK YAMATU TUALANG	Rp1.301.832	Rp656.397	Rp1.432.607	C2
4	SMPN 39 SIAK	Rp328.696	Rp316.922	Rp459.457	C2
5	MUSYALLA MUHAMMADIYA	Rp122.887	Rp768.470	Rp7.902	C3
6	MUSHALLA ISTIQOMAH	Rp116.397	Rp761.981	Rp14.386	C3
7	SMAN 1 SUNGAI MANDAU	Rp645.614	Rp0	Rp776.366	C2
8	NIRWAN 1	Rp101.333	Rp746.920	Rp29.446	C3
9	SD NEGERI 04 PERAWANG BAR	Rp285.646	Rp360.100	Rp416.422	C1
10	SLTP 4 MINAS.	Rp593.951	Rp53.096	Rp724.727	C2
11	SD 020	Rp92.665	Rp738.254	Rp38.114	C3
12	MUSALLA DARUSADAH	Rp103.408	Rp542.231	Rp234.184	C1
13	MESJID DARUSSALAM	Rp115.123	Rp530.503	Rp245.896	C1
14	MULL.DT.MARAJO	Rp78.827	Rp724.429	Rp51.950	C3
15	MUSH.DARUSALAM	Rp83.975	Rp729.576	Rp46.802	C3
16	SD NO.056	Rp64.572	Rp710.177	Rp66.207	C1
17	MUSH TANAH LAPANG	Rp113.675	Rp759.269	Rp17.102	C3
18	MUSH. AMANAH	Rp130.777	Rp776.366	Rp0	C3
19	M.NURUL IKLAS	Rp123.905	Rp769.497	Rp6.871	C3
20	MUSH. M. R A S U L	Rp97.439	Rp743.037	Rp33.338	C3
21	MUSH ISTIQOMAH	Rp74.670	Rp720.224	Rp56.145	C3
22	MUSALLA NURUL FALAH	Rp131.318	Rp776.907	Rp541	C3
23	MESJID DAL FALAH	Rp96.229	Rp741.817	Rp34.550	C3

24	SD. 047	Rp116.722	Rp762.306	Rp14.061	C3
25
60	SMPS LPM	Rp78.835	Rp724.418	Rp51.949	C3

Setiap kolom pada matrik menunjukkan nilai jarak data terhadap pusat *cluster*. Baris pertama pada matrik menunjukkan nilai jarak data terhadap titik pusat *cluster* pertama, dan baris kedua pada matriks menunjukkan nilai jarak data terhadap titik pusat *cluster* kedua begitu juga dengan baris ketiga.

c. Pengelompokan data

Jarak dari hasil perhitungan yang sudah dilakukan, akan dibandingkan dan dipilih jarak terdekat atau nilai terkecil dari ketiga hasil perhitungan jarak setiap *cluster*, jarak ini menunjukkan bahwa data tersebut berada dalam satu kelompok dengan *cluster* terdekat.

Tabel 4. 13 Pengelompokan *Cluster* Pertama

No.	C1	C2	C3
1			*
2	*		
3		*	
4		*	
5			*
6			*
7		*	
8			*
9	*		
10		*	
11			*
12	*		
13	*		
14			*
15			*

16	*		
17			*
18			*
19			*
20			*
21			*
22			*
23			*
24			*
25	.	.	.
60			*

d. Perhitungan titik *centroid* baru

Hitung titik *centroid* baru berdasarkan data anggota kelompok tiap-tiap *cluster*.

Tabel 4. 14 Pusat *Cluster* Baru

	DAYA	TAGIHAN	PEMKWH	JAMNYALA
C1	1623	212614	313	260
C2	6100	791381	782	254
C3	864	28862	63	90

Iterasi ke-2 :

Ulangi langkah kedua (b) hingga data tidak mengalami perubahan. Berikut hasil perhitungan jarak iterasi kedua :

Tabel 4. 15 Perhitungan Jarak Iterasi-2

No.	Nama	C1	C2	C3	Cluster
1	TK AR RAHMAN	Rp185.356	Rp764.138	Rp1.664	C2
2	MESJID NURUL IMAN	Rp125.388	Rp704.171	Rp58.366	C2
3	SMK YAMATU TUALANG	Rp1.224.900	Rp646.115	Rp1.408.653	C2
4	SMPN 39 SIAK	Rp251.770	Rp327.075	Rp435.502	C1

5	MUSYALLA MUHAMMADIYA	Rp199.816	Rp778.599	Rp16.063	C2
6	MUSHALLA ISTIQOMAH	Rp193.326	Rp772.110	Rp9.573	C2
7	SMAN 1 SUNGAI MANDAU	Rp568.691	Rp14.618	Rp752.411	C2
8	NIRWAN 1	Rp178.263	Rp757.046	Rp5.492	C2
9	SD NEGERI 04 PERAWANG BAR	Rp208.718	Rp370.081	Rp392.470	C2
10	SLTP 4 MINAS.	Rp517.020	Rp61.805	Rp700.774	C2
11	SD 020	Rp169.594	Rp748.378	Rp14.160	C2
12	MUSALLA DARUSADAH	Rp26.480	Rp552.310	Rp210.231	C2
13	MESJID DARUSSALAM	Rp38.191	Rp540.595	Rp221.943	C1
14	MULL.DT.MARAJO	Rp155.759	Rp734.544	Rp28.002	C2
15	MUSH.DARUSALAM	Rp160.907	Rp739.692	Rp22.855	C3
16	SD NO.056	Rp141.504	Rp720.288	Rp42.257	C2
17	MUSH TANAH LAPANG	Rp190.606	Rp769.391	Rp6.864	C2
18	MUSH. AMANAH	Rp207.707	Rp786.491	Rp23.955	C3
19	M.NURUL IKLAS	Rp200.836	Rp779.621	Rp17.086	C1
20	MUSH. M. R A S U L	Rp174.371	Rp753.155	Rp9.396	C2
21	MUSH ISTIQOMAH	Rp151.587	Rp730.364	Rp32.194	C2
22	MUSALLA NURUL FALAH	Rp208.248	Rp787.032	Rp24.496	C2
23	MESJID DAL FALAH	Rp173.158	Rp751.942	Rp10.596	C2
24	SD. 047	Rp193.651	Rp772.435	Rp9.898	C2
25

60	SMPS LPM	Rp155.762	Rp734.545	Rp27.994	C2

Langkah berikutnya, melakukan perbandingan perbandingan hasil *cluster* pada iterasi pertama dengan kedua. Jika nilai anggota *cluster* memiliki nilai yang konvergen/sama maka proses iterasi berhenti. Iterasi dihentikan dan hasil *cluster* tersebut merupakan hasil *cluster* yang sudah valid.

Tabel 4. 16 Perbandingan Iterasi

No.	Iterasi 1	Iterasi 2
1	C3	C3
2	C1	C3
3	C2	C2
4	C2	C1
5	C3	C3
6	C3	C3
7	C2	C2
8	C3	C3
9	C1	C1
10	C2	C2
11	C3	C3
12	C1	C1
13	C1	C1
14	C3	C3
15	C3	C3
16	C1	C3
17	C3	C3
18	C3	C3
19	C3	C3
20	C3	C3
21	C3	C3
22	C3	C3
23	C3	C3
24	C3	C3
25	.	.
	.	.
60	C3	C3

Kesimpulannya, jika nilai anggota *cluster* sebelumnya memiliki nilai yang konvergen dengan nilai anggota *cluster* setelahnya maka proses iterasi dihentikan. Jika nilai anggota *cluster* belum konvergen maka ulangi langkah sebelumnya (d) untuk menghitung pusat *cluster* baru.

Sementara hasil dari sistem menunjukkan hasil *cluster* yang sama dari hitungan manual. Berikut ini merupakan hasil sistem pada iterasi terakhir.

No.	ID	Nama	C1	C2	C3	Minimum	Cluster
1	181500144383	TK AR RAI IMAN	Rp321,202.82	Rp816,647.50	Rp25,663.69	Rp25,663.69	c3
2	181500140176	MESJID NURUL IMAN	Rp261,235.63	Rp758.680.41	Rp34,310.04	Rp34,310.04	c3
3	181500060998	SMK YAMTU TUALANG	Rp1,089,051.20	Rp591.607.04	Rp1,384,597.07	Rp591,607.04	c2
4	181500153810	SMPN 39 SIAR	Rp115,553.25	Rp381,587.11	Rp411,446.98	Rp115,553.25	c1
5	181500064479	MUSYAWADA MUHAMMADIYA	Rp335,663.76	Rp833,108.54	Rp40,119.01	Rp40,119.01	c3
6	181500093871	MUSHALLA ISTIQOMAH	Rp329,173.65	Rp826,618.63	Rp33,628.95	Rp33,628.95	c3
7	181500248747	SMAN 1 SUNGAI MANDA	Rp432,867.57	Rp65,655.23	Rp726,360.12	Rp65,655.23	c2
8	181500115532	NIRWAN 1	Rp314,110.20	Rp811,554.95	Rp18,565.49	Rp18,565.49	c3
9	181500039416	SD NEGERI 04 PERAWANG BAR	Rp72,892.24	Rp424,585.82	Rp368,413.89	Rp72,892.24	c1

Gambar 4. 9 Hasil Perhitungan Sistem Tarif Sosial

Dari hasil *training* diatas dari 60 pelanggan terdapat 12 pelanggan di *cluster* 1, 6 pelanggan di *cluster* 2 dan 42 pelanggan di *cluster* 3. Hasil *cluster* tersebut akan diuji performanya dengan menggunakan RapidMiner dengan metode perhitungan Davies Bouldin.

Dari metode perhitungan Davies Bouldin menghasilkan 0,117 yang artinya 11,7% kemiripan data setiap *cluster*.

Selanjutnya, pengujian hasil dari sistem akan dibandingkan dengan hasil manual dengan menggunakan 25 data *testing* pada tarif sosial. Berikut hasil perbandingan pengujian dari sistem dengan manual.

No.	ID	Nama	C1	C2	C3	Minimum	Manual	Cluster
1	181203820472	FUSKESMAS PFMBANTU MENTUL	Rp319,450.55	Rp1,830,293.51	Rp29,006.89	Rp29,006.89	C3	c3
2	181203842011	MDA AL MIRAROKAH	Rp144,399.29	Rp1,855,211.21	Rp4,059.08	Rp4,059.08	C3	c3
3	181203842506	MESJID AL- FURQON	Rp379,245.03	Rp1,890,078.81	Rp30,789.66	Rp30,789.66	C3	c3
4	181203842530	SMPN 3	Rp151,551.87	Rp1,359,281.83	Rp500,005.61	Rp151,551.87	C2	c1
5	181203844654	MUSHALLA AL_JIHAT	Rp387,024.99	Rp1,898,757.72	Rp39,468.71	Rp39,468.71	C3	c3
6	181203848170	SMKN	Rp1,510,833.03	Rp0.00	Rp1,859,289.45	Rp0.00	C2	c2
7	181203856951	SMA N. 1 GN. SAHILAN	Rp372,118.08	Rp1,882,950.89	Rp23,661.62	Rp23,661.62	C3	c3
8	181203865584	SD 025 GUNING	Rp354,248.21	Rp1,865,131.09	Rp5,841.73	Rp5,841.73	C3	c3

Gambar 4. 10 Hasil *Testing* Tarif Sosial

Dari 25 data *testing* pada tarif sosial, hasilnya mendapatkan perbedaan 1 data. Sehingga perbandingannya adalah $22/25 * 100\% = 88\%$ artinya sistem ini dapat digunakan dengan baik.

4.2.1.3 Tarif Rumah Tangga

Tabel 4. 17 Data *Training* Tarif Rumah Tangga

No.	Daya	Rptag	Pekerjaan	Pghsln	kWh	Jam
1	1300	82636	Wiraswasta	3500000	49	38
2	4400	285530	Karyawan Swasta	3600000	183	42
3	1300	109652	Wiraswasta	2000000	69	53
4	900	340569	Wiraswasta	3000000	229	254
5	1300	286051	Wiraswasta	2300000	180	138
6	1300	772336	Karyawan Swasta	4500000	486	374
7	1300	274926	Buruh	3300000	173	133
8	1300	807299	Karyawan Swasta	3700000	508	391
9	1300	236786	Karyawan Swasta	4000000	149	115
10	1300	549853	PNS	4415000	346	266
11	1300	338493	Buruh	3500000	213	164
12	1300	619776	Karyawan Swasta	2800000	390	300
13	1300	201825	Karyawan Swasta	3300000	127	98
14	900	267255	Karyawan Swasta	3500000	631	701
15	900	83078	Wiraswasta	2000000	180	200
16	900	178464	Buruh	3100000	120	133

17	1300	120777	Karyawan Swasta	2500000	76	58
18	1300	109652	Buruh	3200000	69	53
19	900	144258	Wiraswasta	2400000	97	108
20	1300	257445	Karyawan Swasta	3800000	162	125
21	1300	219306	Karyawan Swasta	4100000	138	106
22	1300	82636	Buruh	3800000	50	38
23	1300	147793	PNS	3300000	93	72
24	1300	378223	Karyawan Swasta	4300000	238	183
25
150	900	230516	Wiraswasta	3000000	155	172

Karena atribut pekerjaan bukan merupakan data nominal, maka data dari atribut pekerjaan akan dirubah menjadi bentuk nominal dengan inisialisasi berdasarkan frekuensi data. Berikut inialisasi data desa:

Tabel 4. 18 Inialisasi Pekerjaan

DESA	FREKUENSI	INISIAL
Karyawan Swasta	69	1
Wiraswasta	42	2
Buruh	34	3
PNS	5	4

Iterasi Ke-1 :

- a. Menentukan pusat awal *Cluster*

Nilai Random awal :

Tabel 4. 19 Pusat *Cluster* Awal

C	Daya	Rptag	pkrgan	pghsln	kWh	JNyal
C1	900	144258	2	2400000	97	108
C2	900	391134	2	3200000	263	292
C3	1300	462449	1	4000000	291	224

b. Perhitungan jarak pusat *cluster*

Untuk mengukur jarak antara data dengan pusat *cluster* digunakan *Euclidean distance* dengan rumus sebagai berikut :

Berikut contoh perhitungan jarak dari data ke-1 terhadap pusat *cluster* :

$$d(x_1, c_1) = \sqrt{(900 - 1300)^2 + (144258 - 82636)^2 + (2 - 2)^2 + (240000 - 350000)^2 + (97 - 49)^2 + (108 - 38)^2}$$

$$= 1101725$$

$$d(x_1, c_2) = \sqrt{(900 - 1300)^2 + (391134 - 82636)^2 + (2 - 2)^2 + (320000 - 350000)^2 + (263 - 49)^2 + (292 - 38)^2}$$

$$= 430315$$

$$d(x_1, c_3) = \sqrt{(1300 - 1300)^2 + (462449 - 82636)^2 + (1 - 2)^2 + (400000 - 350000)^2 + (291 - 49)^2 + (224 - 38)^2}$$

$$= 627900$$

Dan seterusnya dilanjutkan untuk data ke 2, 3,..n. Kemudian akan didapatkan matrik jarak untuk iterasi ke-1.

Tabel 4. 20 Perhitungan Jarak Iterasi-1

No.	Nama	C1	C2	C3	Cluster
1	IR. BISNER V SINAGA	Rp1.101.72 5	Rp430.315	Rp627.900	C2
2	SUHERI ROBEL	Rp1.208.29 2	Rp413.720	Rp437.390	C2
3	S.MAKMUR	Rp401.494	Rp1.232.57 1	Rp2.030.87 8	C1
4	KAMARIAH.	Rp631.299	Rp206.293	Rp1.007.40 0	C2
5	BUSTAMI	Rp173.509	Rp906.114	Rp1.709.12 7	C1
6	SUMANTO	Rp2.191.91 3	Rp1.354.73 8	Rp588.243	C3

7	SYAFRIR	Rp909.436	Rp153.312	Rp724.683	C2
8	KAIDIR	Rp1.459.32 3	Rp650.533	Rp457.079	C3
9	HENDRIK N	Rp1.602.67 3	Rp814.754	Rp225.663	C3
10	P.SILALAH	Rp2.055.41 5	Rp1.225.32 3	Rp424.104	C3
11	AHIYEN.	Rp1.117.01 7	Rp304.584	Rp515.136	C2
12	ZULKIFLI	Rp621.384	Rp460.736	Rp1.210.26 9	C2
13	BUSTAMAM	Rp901.839	Rp214.098	Rp746.944	C2
14	RUSTAM YAHYA	Rp1.106.85 5	Rp324.571	Rp536.751	C2
15	SULAIMAN	Rp404.652	Rp1.238.91 0	Rp2.035.66 3	C1
16	KAMARUDDIN. S	Rp700.835	Rp235.008	Rp943.741	C2
17	SYAFRILHARY	Rp102.721	Rp750.395	Rp1.538.42 1	C1
18	JON RIZAL	Rp800.748	Rp281.482	Rp874.337	C2
19	KUSMAN	Rp0	Rp837.226	Rp1.631.33 2	C1
20	CANDRA	Rp1.404.56 8	Rp614.714	Rp286.403	C3
21	AGUS SALIM	Rp1.701.65 6	Rp916.256	Rp262.904	C3
22	RISNAWATI	Rp1.401.35 6	Rp674.664	Rp429.253	C3
23	AEK NABARA	Rp900.007	Rp263.088	Rp767.469	C2
24	GUSRION	Rp1.914.35 1	Rp1.100.07 6	Rp311.599	C3
25
15 0	EFFENDI C 9	Rp606.169	Rp256.512	Rp1.026.54 4	C2

c. Pengelompokan data

Jarak dari hasil perhitungan yang sudah dilakukan, akan dibandingkan dan dipilih jarak terdekat atau nilai terkecil dari ketiga hasil perhitungan jarak setiap

cluster, jarak ini menunjukkan bahwa data tersebut berada dalam satu kelompok dengan *cluster* terdekat.

Tabel 4. 21 Pengelompokan *Cluster* Iterasi-1

No.	C1	C2	C3
1		*	
2		*	
3	*		
4		*	
5	*		
6			*
7		*	
8			*
9			*
10			*
11		*	
12		*	
13		*	
14		*	
15	*		
16		*	
17	*		
18		*	
19	*		
20			*
21			*
22			*
23		*	
24			*
25	.	.	.
	.	.	.
	.	.	.
150		*	

d. Perhitungan titik *centroid* baru

Hitung titik *centroid* baru berdasarkan data anggota kelompok tiap-tiap *cluster*.

Tabel 4. 22 Pusat *Cluster* Baru

C	Desa	Daya	Rptag	pkrgan	pghsln	kWh	JNyala
C1	3	1085	253329	2	2415385	173	161
C2	3	1246	310931	2	3267164	211	179
C3	3	1374	431006	2	4464500	275	203

Iterasi ke-2 :

Ulangi langkah kedua (b) hingga data tidak mengalami perubahan. Berikut hasil perhitungan jarak iterasi kedua :

Tabel 4. 23 Perhitungan Jarak Iterasi-2

No.	Nama	C1	C2	C3	Cluster
1	IR. BISNER V SINAGA	Rp1.097.965	Rp326.085	Rp1.025.486	C2
2	SUHERI ROBEL	Rp1.185.058	Rp333.819	Rp876.660	C2
3	S.MAKMUR	Rp439.531	Rp1.283.050	Rp2.485.363	C1
4	KAMARIAH.	Rp591.089	Rp268.803	Rp1.467.290	C2
5	BUSTAMI	Rp119.935	Rp967.484	Rp2.169.348	C1
6	SUMANTO	Rp2.148.253	Rp1.316.351	Rp343.171	C3
7	SYAFRIR	Rp884.879	Rp48.730	Rp1.174.913	C2
8	KAIDIR	Rp1.398.971	Rp658.580	Rp852.090	C2
9	HENDRIK N	Rp1.584.702	Rp736.577	Rp503.470	C3
10	P.SILALAH	Rp2.021.482	Rp1.172.438	Rp128.743	C3
11	AHIYEN.	Rp1.087.954	Rp234.461	Rp968.927	C2
12	ZULKIFLI	Rp531.237	Rp560.025	Rp1.675.170	C1
13	BUSTAMAM	Rp886.113	Rp113.940	Rp1.186.838	C2
14	RUSTAM YAHYA	Rp1.084.705	Rp236.898	Rp978.302	C2
15	SULAIMAN	Rp448.921	Rp1.287.487	Rp2.488.938	C1
16	KAMARUDDIN. S	Rp688.697	Rp213.287	Rp1.387.674	C2
17	SYAFRILHARY	Rp157.257	Rp790.379	Rp1.988.845	C1
18	JON RIZAL	Rp797.662	Rp212.189	Rp1.304.695	C2
19	KUSMAN	Rp110.151	Rp883.037	Rp2.084.319	C1
20	CANDRA	Rp1.384.622	Rp535.514	Rp686.792	C2
21	AGUS SALIM	Rp1.684.959	Rp837.861	Rp421.518	C3
22	RISNAWATI	Rp1.395.097	Rp579.683	Rp750.281	C2
23	AEK NABARA	Rp890.888	Rp166.410	Rp1.198.445	C2

24	GUSRION	Rp1.888.749	Rp1.035.026	Rp172.761	C3
25
150	EFFENDI C 9	Rp585.060	Rp279.004	Rp1.478.160	C2

Langkah berikutnya, melakukan perbandingan perbandingan hasil *cluster* pada iterasi pertama dengan kedua. Jika nilai anggota *cluster* memiliki nilai yang konvergen/sama maka proses iterasi berhenti. Iterasi dihentikan dan hasil *cluster* tersebut merupakan hasil *cluster* yang sudah valid.

Tabel 4. 24 Perbandingan Iterasi

No.	Iterasi 1	Iterasi 2
1	C2	C2
2	C2	C2
3	C1	C1
4	C2	C2
5	C1	C1
6	C3	C3
7	C2	C2
8	C3	C2
9	C3	C3
10	C3	C3
11	C2	C2
12	C2	C1
13	C2	C2
14	C2	C2
15	C1	C1
16	C2	C2
17	C1	C1
18	C2	C2
19	C1	C1
20	C3	C2
21	C3	C3
22	C3	C2
23	C2	C2
24	C3	C3
25	.	.
150	C2	C2

Kesimpulannya, jika nilai anggota *cluster* sebelumnya memiliki nilai yang konvergen dengan nilai anggota *cluster* setelahnya maka proses iterasi dihentikan. Jika nilai anggota *cluster* belum konvergen maka ulangi langkah sebelumnya (d) untuk menghitung pusat *cluster* baru.

Sementara hasil dari sistem menunjukkan hasil *cluster* yang sama dari hitungan manual. Berikut ini merupakan hasil sistem pada iterasi terakhir.

No.	ID	Nama	C1	C2	C3	Minimum	Cluster
1	181500145302	IR. RISNERY SINAGA	Rp1,036,681.97	Rp236,281.78	Rp1,544,350.65	Rp236,281.78	c2
2	181500180889	SUHERI ROBEI	Rp1,112,679.44	Rp115,228.58	Rp1,403,375.14	Rp115,228.58	c2
3	181500054638	S. MAKMUR	Rp525,272.83	Rp1,504,285.04	Rp3,012,287.31	Rp525,272.83	c1
4	181500057496	KAMARIAH	Rp513,716.21	Rp490,179.28	Rp1,993,275.13	Rp490,179.28	c2
5	181500055299	RIJSTAMI	Rp188,479.59	Rp1,190,138.43	Rp2,695,394.46	Rp188,479.59	c1
6	181500055510	SUMANTO	Rp2,055,994.61	Rp1,107,482.38	Rp557,893.02	Rp557,893.02	c3
7	181500055294	SYAFRIR	Rp813,065.35	Rp194,674.33	Rp1,701,837.56	Rp194,674.33	c2
8	181500055678	KA DIR	Rp1,312,333.12	Rp531,945.74	Rp1,323,070.45	Rp531,945.74	c2
9	181500057211	HENDRIK N	Rp1,514,048.36	Rp516,840.56	Rp1,021,383.27	Rp516,840.56	c2

Gambar 4. 11 Hasil Perhitungan Sistem

Dari hasil *training* diatas dari 150 pelanggan terdapat 3 pelanggan di *cluster* 1, 78 pelanggan di *cluster* 2 dan 7 pelanggan di *cluster* 3. Hasil *cluster* tersebut akan diuji performanya dengan menggunakan RapidMiner dengan metode perhitungan Davies Bouldin. Dari metode perhitungan Davies Bouldin menghasilkan 0,125 yang artinya 12,5% kemiripan data setiap *cluster*.

Selanjutnya, pengujian hasil dari sistem akan dibandingkan dengan hasil manual dengan menggunakan 63 data *testing* pada tarif rumah tangga. Berikut hasil perbandingan pengujian dari sistem dengan manual.

No.	ID	Nama	C1	C2	C3	Minimum	Manual	Cluster
1	181203820472	PUSKESMAS PEMBANTU MENTUL	Rp318,450.55	Rp1,830,283.51	Rp29,006.89	Rp29,006.89	C3	c3
2	181203842011	MDA AL MUBAROKAH	Rp344,339.29	Rp1,855,231.21	Rp4,059.08	Rp4,059.08	C3	c3
3	181203842506	MESJID AL-FURQON	Rp379,245.03	Rp1,890,078.81	Rp30,789.66	Rp30,789.66	C3	c3
4	181203842530	SMPN 3	Rp151,551.87	Rp1,359,281.83	Rp500,005.61	Rp151,551.87	C2	c1
5	181203844654	MUSHALLA AL-JIHAT	Rp387,524.99	Rp1,898,757.72	Rp39,468.71	Rp39,468.71	C3	c3
6	181203848470	SMKN	Rp1,510,833.03	Rp0.00	Rp1,859,289.45	Rp0.00	C2	c2
7	181203856951	SMA N. 1 GN. SAHILAN	Rp372,118.08	Rp1,882,950.89	Rp23,661.62	Rp23,661.62	C3	c3
8	181203855584	SD 025 GUNUNG	Rp350,248.21	Rp1,865,131.09	Rp5,841.73	Rp5,841.73	C3	c3

Gambar 4. 12 Hasil *Testing* Tarif Rumah Tangga

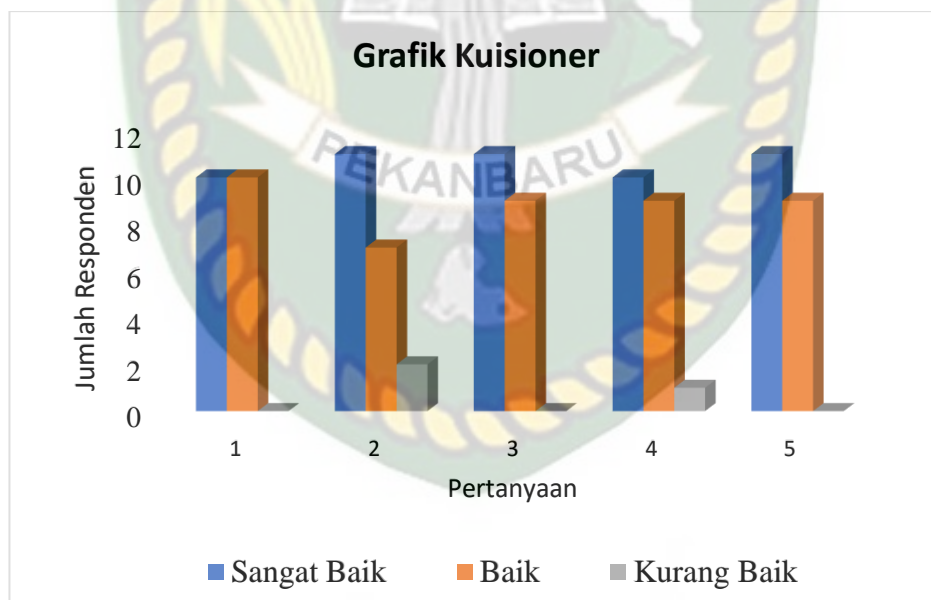
Dari 63 data *testing* pada tarif rumah tangga, hasilnya mendapatkan perbedaan 12 data. Sehingga perbandingannya adalah $51/63 * 100\% = 82,5\%$ artinya sistem ini dapat digunakan dengan baik.

4.3 Implementasi Sistem

Implementasi sistem yang digunakan adalah dengan membuat kuisisioner dengan 5 pertanyaan dan 20 responden umum yang terdiri dari pengguna dan penyedia layanan service sebagai pengguna sistem. Kepada 20 responden diajukan pertanyaan-pertanyaan yang dimaksud adalah sebagai berikut :

1. Apakah aplikasi mudah digunakan (*User Friendly*) ?
2. Sistem ini dapat memberi pengetahuan mengenai karakteristik pengelompokan tunggakan listrik pascabayar ?
3. Bagaimanakah kelengkapan semua fitur dan tampilan aplikasi (*Insert*, *Delete*, dan *Layout*) ?
4. Apakah informasi yang diberikan jelas ?
5. Bagaimanakah tingkat keakuratan informasi ?

Dari 5 pertanyaan diatas, maka diperoleh hasil jawaban atau tanggapan dari responden terhadap kinerja dan tujuan dari sistem ini pada gambar 4.13. Dengan x merupakan kriteria penilaian dan y merupakan total nilai dari pengisi kuisisioner.



Gambar 4. 13 Grafik Hasil Kuisisioner

Keterangan gambar 4.13 :

1. Apakah aplikasi mudah digunakan (*User Friendly*) ? Memiliki nilai Sangat Bagus : 10, Baik : 10, dan Kurang Baik : 0.
2. Sistem ini dapat memberi pengetahuan mengenai karakteristik pengelompokan tunggakan listrik pascabayar ? Memiliki nilai Sangat Bagus : 11, Baik : 7, dan Kurang Baik : 2.
3. Bagaimanakah kelengkapan semua fitur dan tampilan aplikasi (*Insert*, *Delete*, dan *Layout*) ? Memiliki nilai Sangat Bagus : 11, Baik : 9, dan Kurang Baik : 0.
4. Apakah informasi yang diberikan jelas ? Memiliki nilai Sangat Bagus : 10, Baik : 9, dan Kurang Baik : 1.
5. Bagaimanakah tingkat keakuratan informasi ? Memiliki nilai Sangat Bagus : 11, Baik : 9, dan Kurang Baik : 0.

4.3.1 Kesimpulan Implementasi Sistem

Dengan menggunakan *skala likert* maka dapat disimpulkan bahwa sistem *clustering* dengan metode K-Means ini memiliki persentase sebagai berikut :

Tabel 4. 25 Perhitungan Kuisisioner dengan *Skala Likert*

No.	Pertanyaan	Sangat Baik	Baik	Kurang Baik	Skor
1	Pertanyaan 1	10	10	0	$10*3+10*2+0*1 = 50$
2	Pertanyaan 2	11	7	2	$11*3+7*2+2*1 = 49$
3	Pertanyaan 3	11	9	0	$11*3+9*2+0*1 = 51$
4	Pertanyaan 4	10	9	1	$10*3+9*2+1*1 = 52$
5	Pertanyaan 5	11	9	0	$11*3+9*2+0*1 = 51$
Total		53	44	3	253
Persentase	$253/300*100$				84,3%

Dari hasil persentase tabel diatas dengan dasar 5 pertanyaan yang sudah diajukan secara langsung oleh penulis kepada 20 responden, dapat diambil kesimpulan bahwa sistem *clustering* tunggakan rekening listrik pascabayar di PLN ULP Tualang ini memiliki *performance* sangat baik dengan nilai 84,3% sehingga sistem ini layak untuk diimplementasikan.



BAB V

PENUTUP

5.1 Kesimpulan

Berdasarkan hasil analisa data dan perancangan sistem *clustering* tunggakan listrik pascabayar dengan metode K-Means ini dapat disimpulkan bahwa :

1. Sistem ini dapat membantu memberikan hasil perhitungan metode K-Means sesuai dengan perhitungan manual tanpa batas maksimal iterasi dengan nilai pusat *cluster* awal secara random.
2. *Cluster* terbagi menjadi 3 label yaitu tindak lanjuti, dalam monitoring, dan aman. Data yang masuk dalam *cluster* tindak lanjuti akan diberikan sanksi penegasan seperti *invoice*.
3. Sistem ini mengelompokkan data berdasarkan tarifnya masing-masing dengan akurasi tarif bisnis sebesar 95%, tarif sosial 88% dan tarif rumah tangga sebesar 83%.

5.2 Saran

Saran dari penulis untuk sistem *clustering* tunggakan rekening listrik pascabayar dengan metode K-Means ini kedepannya adalah :

1. Penelitian ini sebaiknya menggunakan metode lain seperti metode *KNN* atau metode data *mining* lainnya untuk mengembangkan metode k-means sehingga sistem ini dapat memberikan hasil yang lebih baik.
2. Mengembangkan efek *visual* pada sistem ini dalam bentuk geografis atau bentuk grafik agar lebih mudah dipahami lagi oleh penggunanya.

DAFTAR PUSTAKA

- Dinata, Rozi dkk. (2020). Analisis K-Means Clustering pada Data Sepeda Motor. *Informatics Journal*, 10-17.
- Manik, E. (2018). Data Mining Pengelompokan Pemakaian Listrik Terboros Menggunakan Metode Clustering (Studi Kasus PLN Area Binjai). *Information System Development*, 116-119.
- Primartha, R. (2021). *Algoritma Machine Learning*. Bandung: Informatika Bandung.
- Rahmadani, Nurul dkk. (2020). Implementasi Metode K-Means Clustering Tunggalan Rekening Listrik Pada PT PLN (Persero) Gardu Induk Kisaran. *J-SISCO*, 103-117.
- Rohmawati, Nurul dkk. (2015). Implementasi Algoritma K-Means Dalam Pengklasteran Mahasiswa Pelamar Beasiswa. *JITTER*, 62-68.
- Solichin, A. (2010). *MySQL 5: Dari Pemula Hingga Mahir*. Jakarta: Universitas Budi Luhur.
- Sukrianto, Darmanta. (2017). *Penerapan Teknologi Barcode Pada Pengolahan Data Pembayaran Sumbangan Pembinaan Pendidikan*. Riau: Amik Mahaputra Riau.
- Suyanto, D. (2019). *Data Mining Untuk Klasifikasi dan Klasterisasi Data*. Bandung: Informatika Bandung.

Vulandari, R. T. (2017). *Data Mining Teori dan Aplikasi Rapidminer*. Yogyakarta:

Gava Media.



Dokumen ini adalah Arsip Miik :

Perpustakaan Universitas Islam Riau