

Data Mining Pengelompokan Pembayaran Listrik Menggunakan Metode Clustering: Studi Kasus Kantor Pos Binjai

Adelia Ramadani¹, Achmad Fauzi, S.Kom., M.Kom², Anton Sihombing, SE., MM³

^{1,2,3}Prodi Sistem Informasi STMIK Kaputama,

¹adeliaoppo526@gmail.com, ²fauzyrivai88@gmail.com, ³antonkaputama@gmail.com

ABSTRACT

Electricity is a basic need that greatly influences all community and government activities. Almost all people from the upper to the lower classes, both in the fields of education, government, economics and social, really need a source of electrical energy as a means of supporting activities and productivity. PLN (Persero) is a State-Owned Enterprise (BUMN) that provides services to the community in the provision of electricity. The K-Means algorithm is one of the clustering technique algorithms that starts with random selection, which is the number of clusters that you want to form from the data to be clustered, namely electricity payments. The system made from k-means displays the results of the clustering of electricity payments, namely the pattern of electricity payments whose clusters are up, fixed and down. By utilizing a number of data owned by the agency, it can be grouped using data mining technology. The use of data mining techniques in classifying electricity payments is expected to be useful to simplify the process of searching for system data, which previously was still manual

Keywords : data mining, k-means, electricity.

ABSTRAK

Listrik merupakan kebutuhan mendasar yang sangat berpengaruh terhadap seluruh aktivitas masyarakat maupun pemerintah. Hampir seluruh masyarakat mulai dari kalangan atas sampai kalangan bawah, baik dalam bidang pendidikan, pemerintah, ekonomi maupun sosial sangat membutuhkan sumber energi listrik sebagai sarana penunjang dalam melakukan aktivitas maupun produktivitas. PLN (Persero) merupakan Badan Usaha Milik Negara (BUMN) yang memberikan pelayanan kepada masyarakat dalam penyediaan listrik. Algoritma K-Means merupakan salah satu algoritma teknik clustering yang dimulai dengan pemilihan secara acak, yang merupakan banyaknya cluster yang ingin dibentuk dari data yang akan dicluster yaitu pembayaran listrik. sistem yang dibuat dari k-means menampilkan hasil clustering pembayaran listrik, yaitu pola dari pembayaran listrik yang cluster nya naik, tetap dan turun. Dengan memanfaatkan sejumlah data yang dimiliki pihak instansi, maka dapat dikelompokkan dengan menggunakan teknologi data mining. Penggunaan teknik data mining dalam pengelompokan pembayaran listrik diharapkan dapat berguna untuk mempermudah proses dalam pencarian data secara sistem, yang sebelumnya masih bersifat manual

Kata Kunci : data mining, k-means, listrik.

PENDAHULUAN

Listrik merupakan kebutuhan mendasar yang sangat berpengaruh terhadap seluruh aktivitas masyarakat maupun pemerintah. Hampir seluruh masyarakat mulai dari kalangan atas sampai kalangan bawah, baik dalam bidang pendidikan, pemerintah, ekonomi maupun sosial sangat membutuhkan sumber energi listrik sebagai sarana penunjang dalam melakukan aktivitas maupun produktivitas. Ketergantungan manusia akan listrik terjadi karena listrik merupakan satu-satunya energi yang mampu

menjalankan segala bentuk aktivitas operasional seperti, pengelolaan, pengendalian sarana komunikasi dan informasi di instansi. Sehingga tidak dapat dipungkiri bahwa hampir seluruh aktivitas kehidupan manusia tidak dapat terlepas dari penggunaan tenaga listrik.

PLN (Persero) merupakan Badan Usaha Milik Negara (BUMN) yang memberikan pelayanan kepada masyarakat dalam penyediaan yang berhubungan dengan penjualan listrik di Indonesia guna memenuhi kebutuhan listrik masyarakat Indonesia. serta diberi hak untuk memasok listrik kepada masyarakat, mengatur distribusi dan sumber pemasoknya.

Layanan listrik yang diberikan oleh PLN (Persero) adalah listrik Prabayar dan listrik Pascabayar. Listrik pascabayar merupakan tenaga listrik yang dibayarkan setelah pemakaian oleh konsumen. Sedangkan prabayar merupakan tenaga listrik yang dibayarkan sebelum pemakaian oleh konsumen. Pelayanan jasa yang beragam diberikan oleh pihak PLN (Persero) telah membuat peluang kepada konsumen untuk memilih pelayanan mana yang diinginkan dengan pilihan jenis dan kualitas jasa yang sesuai dengan kebutuhan.

Biasanya masyarakat menggunakan listrik terlebih dahulu dan membayar tagihannya pada bulan berikutnya. sistem perhitungan yang masih manual ini sering menimbulkan masalah seperti seringnya masyarakat tidak disiplin dalam membayar tagihan rekening listrik tiap bulan, kesalahan dalam mencatat meteran, dan pencurian listrik. Sehingga pihak instansi mengeluarkan dan menerapkan sistem listrik prabayar. Listrik prabayar adalah sistem listrik yang menggunakan KWH meter digital, dimana pelanggan harus membeli voucher khusus untuk dapat menggunakan listrik.

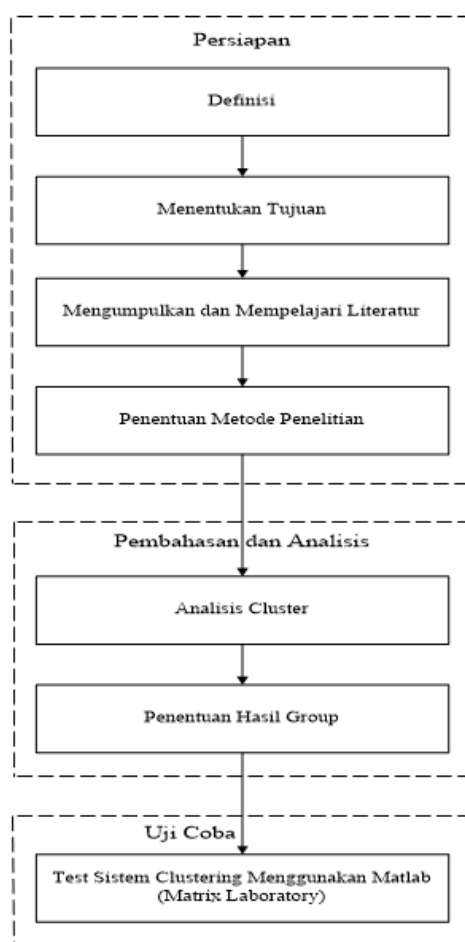
Algoritma K-Means merupakan salah satu algoritma teknik clustering yang dimulai dengan pemilihan secara acak, yang merupakan banyaknya cluster yang ingin dibentuk dari data yang akan di cluster yaitu pembayaran listrik. sistem yang dibuat dari k-means menampilkan hasil clustering pembayaran listrik, yaitu pola dari pembayaran listrik yang cluster nya naik, tetap dan turun. Dengan memanfaatkan sejumlah data yang dimiliki pihak instansi, maka dapat dikelompokkan dengan menggunakan teknologi data mining. Data mining dapat membantu pihak instansi dalam menggali pengetahuan baru dengan cara memproses data pembayaran listrik yang sudah ada. Penggunaan teknik data mining dalam pengelompokan pembayaran listrik diharapkan dapat berguna untuk mempermudah proses dalam pencarian data secara sistem, yang sebelumnya masih bersifat manual.

Metode Penelitian

Metodologi Penelitian ini dilakukan untuk mencari sesuatu sistematis dengan menggunakan metode ilmiah serta sumber yang berlaku. Dengan adanya proses ini dapat memberikan hasil penelitian yang baik dan tepat. pengumpulan data yang digunakan dalam penelitian ini adalah:

- 1) Uraian Kegiatan Penelitian

Penulis melakukan penelitian dengan menyusun langkah-langkah dengan terstruktur agar hasil dari penelitian lebih baik. Berikut ini adalah kerangka kerja uraian kegiatan penelitian mulai dari awal hingga hasil:



Gambar 1. Kerangka Kerja Uraian Kegiatan Penelitian

- a. Definisi
Kegiatan pada tahap ini adalah mendefinisikan pokok permasalahan dimulai dari mencari isu-isu penting seputar algoritma *K-Means*.
- b. Menentukan Tujuan
Setelah mendefinisikan pokok permasalahan, langkah berikut memahami masalah dan menganalisis permasalahan dengan tujuan memperjelas dan menentukan permasalahan sehingga didapat tujuan.
- c. Mengumpulkan dan Mempelajari Literatur
Metode pengumpulan data yang digunakan pada tahap ini adalah Penelitian Kepustakaan (*Library Research*) yaitu dengan mengumpulkan bahan-bahan bacaan berupa karya ilmiah yang bersifat teoritis sebagai landasan atau kerangka teoritis

sesuai dengan judul dan masalah yang dikemukakan dan studi pustaka dapat dihasilkan dari *browsing* di situs-situs internet.

d. Penentuan Metode Penelitian

Agar hasil dapat tercapai dengan baik maka diterapkan metode analisa deskriptif, yaitu data dikumpulkan kemudian disusun, dikelompokkan dan dianalisa sehingga diperoleh gambaran yang jelas pada pokok permasalahan.

e. Analisa *Cluster* (Algoritma *K-Means*)

Setelah data dikumpulkan dan dipelajari langkah berikutnya melakukan analisa *Cluster* (mengelompokkan) data dengan *K-Means* dan variabel yang telah ditentukan daya, total pembayaran dan denda.

f. Penentuan Hasil Group

Pada tahap ini, setelah mengetahui group kemudian di tetapkan maka dapat kita lihat hasilnya. Uji Coba

Tahap ini adalah tahap terakhir dengan menguji penerapan hasil group dari *Cluster* dan di import ke Matlab (*Matrix Laboratory*).

2) Jenis Penelitian

Jenis penelitian yang digunakan dalam penyusunan skripsi ini, dengan metode deskriptif adalah pencarian fakta dengan interpretasi yang tepat. Penelitian deskriptif mempelajari masalah-masalah dalam masyarakat, serta proses-proses yang sedang berlangsung dan pengaruh dari suatu fenomena.

3) Lokasi Penelitian

Lokasi penelitian adalah tempat dimana penelitian akan dilakukan. Tempat penelitian yang akan dilakukan berlokasi di Kantor Pos Binjai.

4) Data dan Sumber Data

a. Data Primer

Data primer adalah data yang dapat diperoleh langsung dari lapangan atau tempat penelitian. Kata-kata dan tindakan merupakan sumber data yang diperoleh dari lapangan dengan mengamati dan mewawancarai. Peneliti menggunakan data ini untuk mendapatkan informasi dari pegawai.

b. Data Sekunder

Data Sekunder adalah data yang didapat dari sumber bacaan dari berbagai sumber lainnya yang terdiri dari, dokumen-dokumen resmi dari berbagai instansi pemerintah. Data sekunder juga dapat berupa hasil studi, hasil wawancara, hasil survey, publikasi dari organisasi, internet dan lain sebagainya. Peneliti menggunakan data sekunder ini untuk memperkuat penemuan dan melengkapi informasi yang telah dikumpulkan melalui wawancara langsung dari pegawai.

5) Teknik Pengumpulan Data

Penelitian dilakukan langsung ke Kantor Pos Binjai untuk mendapatkan data dan sistem pembuatan surat izin praktek yang sedang berjalan. Dimana akan melakukan pengumpulan data dengan dua cara yaitu:

a. Observasi

Observasi yang dilakukan yaitu mengamati dan mengetahui secara langsung jalan sistem yang sedang berjalan saat itu dan proses kerja dan melihat laporan dalam kantor tersebut yang digunakan pada saat ini.

b. Wawancara (*Interview*)

Wawancara yang dilakukan yaitu dengan mengadakan tanya jawab langsung dengan seorang pegawai yang melakukan pengelola data. Untuk mengetahui apa dan bagaimana dari kegiatan.

c. Penelitian Kepustakaan (*Library Research*)

Penulis mengunjungi perpustakaan untuk memperoleh data teoritis yaitu dengan penyesuaian konsep dan teori yang diajarkan di kampus dan praktek di lapangan, maka akan melakukan penelitian ke perpustakaan dengan membaca buku, majalah yang mempunyai hubungannya dengan topik skripsi yang dilakukan.

Data Mining

Data mining berisi pencarian pola yang diinginkan dalam database besar untuk membantu pengambilan keputusan di waktu akan datang. Pola-pola ini dikenal oleh perangkat tertentu yang dapat memberikan suatu analisa data yang berguna dan kemudian dapat dipelajari dengan lebih teliti, yang mungkin saja menggunakan perangkat pendukung keputusan yang lain.

Menurut Prasetyo (2012, h.2) menyatakan bahwa: “*Data mining* merupakan proses untuk mendapatkan informasi yang berguna dari gudang basis data yang besar”. *Data mining* juga dapat diartikan sebagai pengekstrakan informasi yang baru diambil dari bongkahan data besar yang membantu dalam pengambilan keputusan.

Menurut Menurut Hermawati (2013, h. 3) menyatakan bahwa: “*Data mining* adalah proses yang mempekerjakan satu atau lebih teknik pembelajaran komputer (*mechine learning*) untuk menganalisa dan mengekstraksi pengetahuan (*knowledge*) secara otomatis”.

2.1. Algoritma K-Means

Menurut Vlandari (2017, h. 54) *k-means* merupakan algoritma yang menetapkan nilai-nilai *cluster* (*k*) secara *random*, untuk sementara nilai tersebut menjadi pusat dari *cluster* yang biasa disebut *centroid*. Kemudian menghitung jarak setiap data yang ada terhadap masing-masing *centroid* menggunakan rumus $\sqrt{(x - x_c)^2 + (y - y_c)^2}$ hingga ditemukan jarak yang paling dekat dari setiap data dengan *centroid* hingga nilai *centroid* tidak berubah (stabil).

K-Means merupakan salah satu metode pengelompokan data (*nonhirarki*) yang berusaha mempartisi data yang ada kedalam bentuk dua atau lebih kelompok. Adapun tujuan pengelompokan data ini adalah untuk meminimalkan fungsi objektif yang diset dalam proses pengelompokan. Pengelompokan data dengan menggunakan metode *K-Means* ini secara umum dilakukan dengan algoritma dasar sebagai berikut:

- 1) Tentukan jumlah kelompok (*cluster*).
- 2) Alokasikan data ke dalam cluster secara acak (*random*).

- 3) Hitung *centroid*/ rata-rata dari data yang ada di masing-masing *cluster*.
- 4) Alokasikan masing-masing data ke *centroid*/ rata-rata terdekat.

Algoritma *K-Means* memerlukan tiga komponen yaitu:

- 1) Jumlah Kluster
Seperti yang telah dijelaskan sebelumnya, *K-Means* merupakan bagian dari metode non-hirarki sehingga dalam metode ini jumlah *K* harus ditentukan terlebih dahulu. Jumlah kluster *K* dapat ditentukan melalui pendekatan metode hirarki.
- 2) Kluster awal
Kluster awal yang dipilih berkaitan dengan penentuan pusat kluster awal (*centroid* awal).
- 3) Ukuran Jarak
Dalam hal ini, ukuran jarak digunakan untuk menempatkan observasi ke dalam kluster berdasarkan *centroid* terdekat. Ukuran jarak yang digunakan dalam metode *K-Means* adalah *Euclid*.

Adapun algoritma *K-Means* dalam pembentukan kluster sebagai berikut:

- 1) Misalkan diberikan matriks $X = \{X_{ij}\}$ data berukuran $n \times p$ dengan $i=1,2,\dots,n, j=1,2,\dots,p$ dan asumsi jumlah kluster awal *K*
- 2) Tentukan *centroid*
- 3) Hitung jarak setiap objek ke setiap *centroid* dengan menggunakan jarak *Euclid* atau dapat ditulis sebagai berikut:
$$J(x,c) = \sqrt{(x_1 - c_2)^2} \dots \dots \dots (1)$$
- 4) Setiap objek disusun ke *centroid* terdekat dan kumpulkan objek tersebut akan membentuk kluster.
- 5) Tentukan *centroid* baru dari *cluster* yang akan dibentuk, diobjek yang mana *centroid* baru diperoleh dari rata-rata setiap yang terletak pada kluster yang sama.
- 6) Ulangi langkah 3, jika *centroid* awal dan baru tidak sama.

Karakteristik *K-Means* dapat diringkas menjadi seperti berikut:

- 1) *K-Means* merupakan metode pengelompokan yang sederhana dan dapat digunakan dengan mudah.
- 2) Pada jenis set data tertentu, *K-Means* tidak dapat melakukan segmentasi data dengan baik dimana hasil segmentasinya tidak dapat memberikan pola kelompok yang mewakili karakteristik bentuk alami data.
1. *K-Means* bisa mengalami masalah ketika mengelompokkan data yang mengandung *outlier*

Hasil Dan Pembahasan

Penerapan Metode Clustering

Dalam permasalahan metode yang digunakan adalah *clustering*, proses awal yang dilakukan dalam pembentukan *cluster* adalah transformasi data kedalam bentuk numeric dengan kode-kode yang telah ditentukan, tentukan jumlah group (*K*), hitung

centroid, hitung jarak objek ke *centroid* dan kemudian groupkan berdasarkan jarak terdekat, jika tidak ada objek yang pindah group maka iterasi selesai.

Tabel 1. Data Yang akan Diolah

No	Nama	Daya	Total Pembayaran	Denda
1	A	900 VA	210.000	3.000
2	B	450 VA	138.000	5.000
3	C	1.300 VA	525.000	50.000
4	D	900 VA	180.000	3.000
5	E	2.200 VA	642.000	10.000
6	F	1.300 VA	855.000	75.000
7	G	450 VA	377.000	5.000
8	H	900 VA	434.000	50.000
9	I	3.500 VA	1.150.000	100.000
10	J	450 VA	400.000	3.000
11	K	900 VA	482.000	10.000
12	L	2.200 VA	915.000	100.000
13	M	1.300 VA	584.000	5.000
14	N	2.200 VA	761.000	75.000
15	O	450 VA	295.000	3.000
16	P	1.300 VA	600.000	50.000
17	Q	900 VA	453.000	3.000
18	R	1.300 VA	658.000	5.000
19	S	900 VA	320.000	10.000
20	T	3.500 VA	970.000	3.000

1. Inisialisasi

Dari data yang ada maka dapat dilakukan inisialisasi data sesuai dengan kebutuhan variabel sebagai berikut:

a. Inisialisasi Kriteria Daya

Berikut dibawah ini adalah tabel inisialisasi daya yang dapat dilihat pada tabel dibawah ini:

Tabel 2. Inisialisasi Kriteria Daya

Kode	Daya
1	450 VA
2	900 VA
3	1.300 VA
4	2.200 VA
5	> 3.500 VA

El-Mujtama: Jurnal Pengabdian Masyarakat

Vol 2 No 3 (2022) 412 - 428 P-ISSN 2746-9794 E-ISSN 2747-2736

DOI: 10.47467/elmujtama.v2i3.1947

b. Inisialisasi Kriteria Total Pembayaran

Berikut dibawah ini adalah tabel inisialiasi kriteria total pembayaran yang dapat dilihat pada tebal dibawah ini:

Tabel 3. Inisialisasi Kriteria Total Pembayaran

Kode	Total Pembayaran
1	< 200.000
2	200.001 – 400.000
3	400.001 – 600.000
4	600.001 – 800.000
5	800.001 – 1.00.000
6	> 1.000.001

c. Inisialisasi Kriteria Denda

Berikut dibawah ini adalah tabel inisialiasi kriteria denda yang dapat dilihat pada tabel dibawah ini:

Tabel 4. Inisialisasi Kriteria Denda

Kode	Denda
1	3.000 Perbulan
2	5.000 Perbulan
3	10.000 Perbulan
4	50.000 Perbulan
5	75.000 Perbulan
6	100.000 Perbulan

2. Transformasi Data

Agar data diatas dapat diolah dengan menggunakan algoritma *K-Means clustering*, maka data yang berjenis data nonnominal seperti daya, total pembayaran dan denda harus diinisialisasikan terlebih dahulu dalam bentuk angka. Data pembayaran listrik ini dapat dinyatakan dalam suatu variabel-variabel yang independen yaitu daya (X), total pembayaran (Y), dan denda (Z).

Tabel 5 Transformasi Data Berdasarkan Inisialisasi Yang Dibentuk

No	Nama	Daya (X)	Total Pembayaran (Y)	Denda (Z)
1	A	2	2	1
2	B	1	1	2
3	C	3	3	4
4	D	2	1	1
5	E	4	4	3
6	F	3	5	5
7	G	1	2	2
8	H	2	3	4
9	I	5	6	6
10	J	1	2	1
11	K	2	3	3
12	L	4	5	6

13	M	3	3	2
14	N	4	4	5
15	O	1	2	1
16	P	3	3	4
17	Q	2	3	1
18	R	3	4	2
19	S	2	2	3
20	T	5	5	1

3. Perhitungan Data Berdasarkan Algoritma K-Means Clustering

Untuk menentukan group dari satu objek, pertama yang harus dilakukan adalah mengukur jarak *Deuclidean* antara dua titik atau objek atau X, Y dan Z yang didefinisikan sebagai berikut:

$$Deuclidean (X, Y, Z) = \sqrt{\sum(X1 - X2)^2 + (Y1 - Y2)^2 + (Z1 - Z2)^2}$$

Lakukan *cluter* menjadi 3 kelompok (K=3) dan tentukan titik pusat *centroid*. Adapun proses perhitungan clustering seperti dibawah ini:

Iterasi I

Centroid 1 = (2,2,1) diambil dari secara acak dari data 1

Centroid 2 = (1,1,2) diambil dari secara acak dari data 2

Centroid 3 = (3,3,4) diambil dari secara acak dari data 3

Keterangan:

Penentuan *centroid* awal dilakukan secara random.

Penyelesaian:

Bagian A (2,2,1)

K = 3, *centroid*1 = (2,2,1), *centroid*2 = (1,1,2), *centroid* 3 = (3,3,4)

Jarak dari C 1 (X) = $\sqrt{(2 - 2)^2 + (2 - 2)^2 + (1 - 1)^2} = 0$

Jarak dari C 2 (Y) = $\sqrt{(2 - 1)^2 + (2 - 1)^2 + (1 - 2)^2} = 1,73$

Jarak dari C 3 (Z) = $\sqrt{(2 - 3)^2 + (2 - 3)^2 + (1 - 4)^2} = 3,32$

Bagian B (1,1,2)

K = 3, *centroid*1 = (2,2,1), *centroid*2 = (1,1,2), *centroid* 3 = (3,3,4)

Jarak dari C 1 (X) = $\sqrt{(1 - 2)^2 + (1 - 2)^2 + (2 - 1)^2} = 1,73$

Jarak dari C 2 (Y) = $\sqrt{(1 - 1)^2 + (1 - 1)^2 + (2 - 2)^2} = 0$

Jarak dari C 3 (Z) = $\sqrt{(1 - 3)^2 + (1 - 3)^2 + (2 - 4)^2} = 3,46$

Setelah seluruh data dihitung maka didapat hasil penentuan group 1 sebagai berikut:

Tabel 6. Hasil Penentuan Group I

No	Nama	Daya	Total Pembayaran	Denda	Jarak Dari C1	Jarak Dari C2	Jarak Dari C3	Group
		X	Y	Z				
1	A	2	2	1	0	1,73	3,32	1

El-Mujtama: Jurnal Pengabdian Masyarakat

Vol 2 No 3 (2022) 412 - 430 P-ISSN 2746-9794 E-ISSN 2747-2736

DOI: 10.47467/elmujtama.v2i3.1947

2	B	1	1	2	1,73	0	3,46	2
3	C	3	3	4	3,32	3,46	0	3
4	D	2	1	1	1	1,41	3,74	1
5	E	4	4	3	3,46	4,36	1,73	3
6	F	3	5	5	5,10	5,38	2,24	3
7	G	1	2	2	1,41	1	3	2
8	H	2	3	4	3,16	3	1	3
9	I	5	6	6	7,07	7,55	4,12	3
10	J	1	2	1	1	1,41	3,74	1
11	K	2	3	3	2,24	2,45	1,41	3
12	L	4	5	6	6,16	6,40	3	3
13	M	3	3	2	1,73	2,83	2	1
14	N	4	4	5	4,90	5,20	1,73	3
15	O	1	2	1	1	1,41	3,74	1
16	P	3	3	4	3,32	3,46	0	3
17	Q	2	3	1	1	2,45	3,16	1
18	R	3	4	2	2,45	3,61	2,23	3
19	S	2	2	3	2	1,73	1,73	2
20	T	5	5	1	4,24	5,74	4,12	3

Group berdasarkan jarak minimal ke centroid terdekat yaitu:

Jika jarak terpendek berada di C1 maka data di masukan di kelompok 1

Jika jarak terpendek berada di C2 maka data di masukan di kelompok 2

Jika jarak terpendek berada di C3 maka data di masukan di kelompok 3

Grup Lama = {0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0}

Group Baru = {1,2,3,1,3,3,2,3,3,1,3,3,1,3,1,3,2,3}

Terjadi perubahan group, maka dilanjutkan ke iterasi berikutnya.

Untuk group 1 ada 6 data yaitu:

$$C1 = (2+2+1+3+1+2) / 6 = 2,45$$

$$C2 = (2+1+2+3+2+3) / 6 = 2,16$$

$$C3 = (1+1+1+2+1+1) / 6 = 1,16$$

Untuk group 2 ada 3 data yaitu:

$$C1 = (1+1+2) / 3 = 1,33$$

$$C2 = (1+2+2) / 3 = 1,66$$

$$C3 = (2+2+3) / 3 = 2,33$$

Untuk group 3 ada 11 data yaitu:

$$C1 = (3+4+3+2+5+2+4+4+3+3+5) / 11 = 3,45$$

$$C2 = (3+4+5+3+6+3+5+4+3+4+5) / 11 = 4,09$$

$$C3 = (4+3+5+4+6+3+6+5+4+2+1) / 11 = 3,91$$

Iterasi II

Centroid 1 = (2.45, 2.16, 1.16)

Centroid 2 = (1.33, 1.66, 2.33)

Centroid 3 = (3.45, 4.09, 3.91)

Penyelesaian:

Bagian A (2;2;1)

$K = 3$, centroid1 = (2.45;2.16;1.16), centroid2 = (1.33;1.66;2.33), centroid3 = (3.45;4.09;3.91)

Jarak dari C 1 (X) = $\sqrt{(2 - 2,45)^2 + (2 - 2,16)^2 + (1 - 1,16)^2} = 0,49$

Jarak dari C 2 (Y) = $\sqrt{(2 - 1,33)^2 + (2 - 1,66)^2 + (1 - 2,33)^2} = 1,53$

Jarak dari C 3 (Z) = $\sqrt{(2 - 3,45)^2 + (2 - 4,09)^2 + (1 - 3,91)^2} = 3,86$

Bagian B (1;1;2)

$K = 3$, centroid1 = (2.45;2.16;1.16), centroid2 = (1.33;1.66;2.33), centroid3 = (3.45;4.09;3.91)

Jarak dari C 1 (X) = $\sqrt{(1 - 2,45)^2 + (1 - 2,16)^2 + (2 - 1,16)^2} = 2,03$

Jarak dari C 2 (Y) = $\sqrt{(1 - 1,33)^2 + (1 - 1,66)^2 + (2 - 2,33)^2} = 0,81$

Jarak dari C 3 (Z) = $\sqrt{(1 - 3,45)^2 + (1 - 4,09)^2 + (2 - 3,91)^2} = 4,38$

Setelah seluruh data dihitung maka didapat hasil penentuan group II sebagai berikut:

Tabel 7. Hasil Penentuan Group II

No	Nama	Daya	Total Pembayaran	Denda	Jarak Dari C1	Jarak Dari C2	Jarak Dari C3	Group
		X	Y	Z				
1	A	2	2	1	0,49	1,53	3,86	1
2	B	1	1	2	2,03	0,81	4,38	2
3	C	3	3	4	3,01	2,71	1,18	3
4	D	2	1	1	1,25	1,63	4,49	1
5	E	4	4	3	3,03	3,61	1,07	3
6	F	3	5	5	4,81	4,59	1,49	3
7	G	1	2	2	1,68	0,58	3,74	2
No	Nama	Daya	Total Pembayaran	Denda	Jarak Dari C1	Jarak Dari C2	Jarak Dari C3	Group
		X	Y	Z				
8	H	2	3	4	3,00	2,24	1,81	3
9	I	5	6	6	6,68	6,77	3,23	3
10	J	1	2	1	1,47	1,51	4,34	1
11	K	2	3	3	2,07	2,15	2,03	3
12	L	4	5	6	5,82	5,63	2,34	3
13	M	3	3	2	1,36	2,17	2,24	1
14	N	4	4	5	4,53	4,44	1,22	3
15	O	1	2	1	1,47	1,51	4,34	1
16	P	3	3	4	3,01	2,71	1,18	3
17	Q	2	3	1	0,96	2	3,43	1

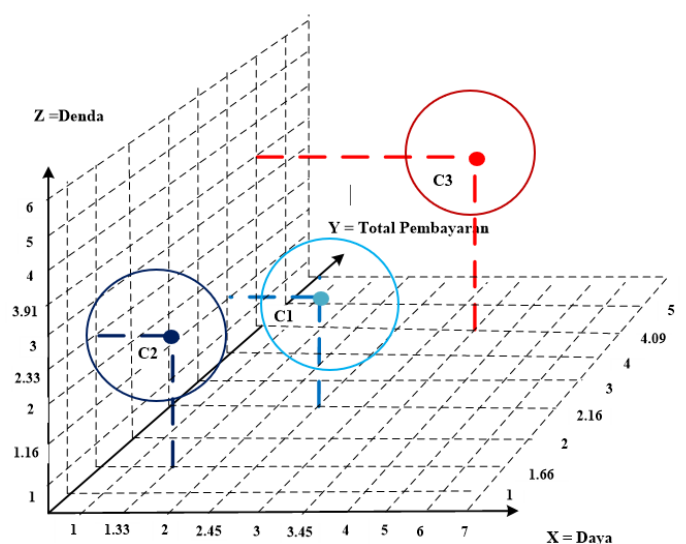
18	R	3	4	2	2,10	2,89	1,96	3
19	S	2	2	3	1,9	1,01	2,70	2
20	T	5	5	1	3,82	5,14	3,42	3

Group berdasarkan jarak minimal ke *centroid* terdekat yaitu:

Group Lama: {1,2,3,1,3,3,2,3,3,1,3,3,1,3,1,3,2,3}

Group Baru: {1,2,3,1,3,3,2,3,3,1,3,3,1,3,1,3,2,3}

Karena pada iterasi ke-1 dan ke-2 *cluster* tidak berubah atau terdapat persamaan, maka perhitungan iterasi dihentikan dan mendapatkan hasil sebagai berikut:



Gambar 2. Grafik *Clustering* Pembayaran Listrik

Keterangan :

- | | X | Y | Z |
|----------------|-------|------|-------|
| ● Centroid 1 : | (2.45 | 2.16 | 1.16) |
| ● Centroid 2 : | (1.33 | 1.66 | 2.33) |
| ● Centroid 3 : | (3.45 | 4.09 | 3.91) |

Penjelasan dari hasil diatas ialah:

Dari 20 data terdapat 3 grup yaitu grup 1 terdapat 6 data, grup 2 terdapat 3 data dan grup 3 terdapat 11 data. Adapun penjelasan terdapat dari 3 grup tersebut sebagai berikut:

1) Cluster 1 Terdapat 6 Data

2.45; 2.16; 1.16;

Berdasarkan perhitungan diatas dapat diketahui bahwasannya pada *cluster* 1 kelompok pembayaran listrik pada grup Daya (X) adalah 900 VA, untuk kelompok Total Pembayaran (Y) ialah 200.001 – 400.000 dan pada grup Denda (Z) yang harus di keluarkan ialah 3.000 Perbulan.

2) Cluster 2 Terdapat 3 Data

1.33; 1.66; 2.33;

Berdasarkan perhitungan diatas dapat diketahui bahwasannya pada *cluster* 1 kelompok pembayaran listrik pada grup Daya (X) adalah 450 VA, untuk kelompok Total Pembayaran (Y) ialah 200.001 – 400.000 dan pada grup Denda (Z) yang harus di keluarkan ialah 5.000 Perbulan.

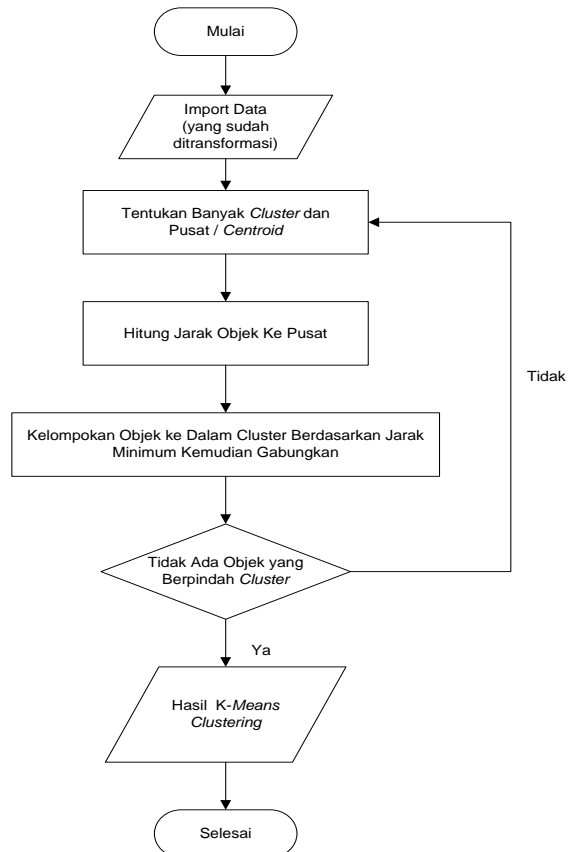
3) Cluster 3 Terdapat 11 Data

3.45; 4.09; 3.91;

Berdasarkan perhitungan diatas dapat diketahui bahwasannya pada *cluster* 1 kelompok pembayaran listrik pada grup Daya (X) adalah 1.300 VA, untuk kelompok Total Pembayaran (Y) ialah 600.001 – 800.000 dan pada grup Denda (Z) yang harus di keluarkan ialah 50.000 Perbulan.

3.1. Flowchart

Adapun rancangan *flowchart* pada proses pengelompokan menggunakan *software* Matlab tentang daya, total pembayaran dan denda dapat digambarkan sebagai berikut:



Gambar 2. Flowchart

Keterangan:

1. Mulai.
2. Ambil data yang akan di proses.
3. Lakukan proses data mining.
4. Kelompokkan objek ke dalam *cluster* berdasarkan jarak terdekat, kemudian gabungkan.
5. Apakah objek ada yang berpindah, jika tidak ada yang berpindah maka lanjut ke langkah selanjutnya dan jika objek ada yang berpindah maka lakukan ke proses data mining lagi.
6. Menampilkan hasil dari *K-Means cluster*.
7. Selesai.

2. Kesimpulan

Berdasarkan hasil penelitian yang telah dilakukan, maka didapat kesimpulan dari penelitian menggunakan algoritma K-Means yaitu sebagai berikut:

1. Penerapan data mining metode clustering dalam pengelompokan pembayaran listrik cukup efektif dalam mengelompokkan pembayaran listrik
2. Dari pengelompokan pembayaran listrik yang terbagi kedalam 3 cluster menghasilkan informasi pada cluster 1 kelompok pembayaran listrik pada grup Daya (X) adalah 900 VA, untuk kelompok Total Pembayaran (Y) ialah 200.001 – 400.000 dan pada grup Denda (Z) yang harus di keluarkan ialah 3.000 Perbulan. Pada cluster 2 kelompok pembayaran listrik pada grup Daya (X) adalah 450 VA, untuk kelompok Total Pembayaran (Y) ialah 200.001 – 400.000 dan pada grup Denda (Z) yang harus di keluarkan ialah 5.000 Perbulan. Pada cluster 3 kelompok pembayaran listrik pada grup Daya (X) adalah 1.300 VA, untuk kelompok Total Pembayaran (Y) ialah 600.001 – 800.000 dan pada grup Denda (Z) yang harus di keluarkan ialah 50.000 Perbulan.

Daftar Pustaka

- Alfatta, H, 2007, *Analisis dan Perancang Sistem Informasi*, Penerbit CV. Andi, Yogyakarta.
- Hermawati, A, F, 2013. *Data Mining*, Penerbit CV. Andi, Yogyakarta.
- Kusrini dan Luthfi, 2009, *Algoritma Data Mining*, Penerbit CV. Andi, Yogyakarta.
- Metisen, M, B. 2015, *Analisis Clustering Menggunakan Metode K-Means Dalam Pengelompokkan Penjualan Produk Pada Swalayan Fadhila*.
- Nugroho, A, 2013, T. *Konsep dan Pemrograman GUI*, Penerbit CV. Andi, Yogyakarta.
- Prasetyo, E, 2012, *Data Mining – Konsep dan Aplikasi Menggunakan Matlab Edisi Pertama*, Yogyakarta, Penerbit Andi, 2012.
- Santosa, B. 2007. *Data Mining Terapan Dengan Matlab*, Yogyakarta.

El-Mujtama: Jurnal Pengabdian Masyarakat

Vol 2 No 3 (2022) 412 - 430 P-ISSN 2746-9794 E-ISSN 2747-2736

DOI: 10.47467/elmujtama.v2i3.1947

Sutrisno, S. 2013. *Penerapan Data Mining Pada Penjualan Menggunakan Metode Clustering Study Kasus PT. Indomarco Palembang.*

Siyamto, Y. 2017, *Pemanfaatan Data Mining Dengan Metode Clustering Untuk Evaluasi Biaya Dokumen Ekspor Di PT Winstar Bata.*