



KILAT

JURNAL KAJIAN ILMU DAN TEKNOLOGI

Rakhmadi Irfansyah Putra

PERANCANGAN APLIKASI PENGEMBALIAN BERKAS TERHAPUS PADA NTFS

Rizqia Cahyaningtyas
Awit Lela Sigi

PERANCANGAN APLIKASI PENGELOLAAN RUMAH TANGGA
LABORATORIUM KOMPUTER STT-PLN

Dian Hartanti

MODEL CLUSTERING MENGGUNAKAN ALGORITMA K-MEANS PADA DATA
KELUHAN PELANGGAN PT PLN (PERSERO)
(STUDI KASUS : PT PLN (PERSERO) DISTRIBUSI JAKARTA DAN TANGERANG)

Abdul Haris; Nina Nirmaya

APLIKASI KONVERSI AKSARA SUNDA KE BAHASA INDONESIA BERBASIS WEB
MENGGUNAKAN PHP MYSQL

Riki Ruli A. Siregar
Alwi Baraqbah

SISTEM KONTROL PADA PENGGUNAAN ENERGI LISTRIK BERBASIS ANDROID
STUDI KASUS : SEKOLAH MENENGAH PERTAMA NEGERI (SMPN) 7 PONTIANAK

Yessy Asri

ANALISA PERBANDINGAN KEPUTUSAN METODE KLASIFIKASI DECISION TREE
DAN NAÏVE BAYES DALAM PENENTUAN DIAGNOSA HIPERTENSI

Marliana Sari

PUSAT INFORMASI KEMAHASISWAAN DENGAN MENGGUNAKAN PHP, MYSQL
DAN METODE MVC

Inge Handriani

FLOWCHART SISTEM PENAGIHAN PADA PERUSAHAAN JASA KONSULTAN

Herman Bedi Agtriadi

RANCANG BANGUN SISTEM MONITORING PADA BANGUNAN PENGENDALI
SEDIMEN BERBASIS ANDROID DI PLTA

Indah Handayasari
Aziz Maulana

DESAIN ALTERNATIF JEMBATAN MENGGUNAKAN PLAT GIRDER
(STUDI KASUS JEMBATAN RSUD KOTA TANGERANG)

Irma Wirantina Kustanrika

PERHITUNGAN SINYAL PADA SIMPANG DENGAN METODE WEBSTER

Mukhlis Akhadi

MEMPRODUKSI BAHAN SEMIKONDUKTOR DI DALAM TERAS REAKTOR NUKLIR

ISSN 2089-1245



9 772089 124519

SEKOLAH TINGGI TEKNIK - PLN (STT-PLN)

KILAT

VOL.4

NO.1

HAL. 1 - 119

APRIL 2015

ISSN 2089 - 1245

MODEL CLUSTERING MENGGUNAKAN ALGORITMA K-MEANS PADA DATA KELUHAN PELANGGAN PT. PLN (STUDI KASUS : PT. PLN (PERSERO) DISTRIBUSI JAKARTA DAN TANGERANG)

Dian Hartanti

Dosen Teknik Informatika STT PLN
(dhianiez.smart9@gmail.com)

Abstrak

Data mining adalah proses mengubah sejumlah besar data sampai menghasilkan keterhubungan antar isi data, proses pengelompokan data (data clustering) menjadi sangat penting karena adanya peningkatan jumlah data dalam format teks yang cukup signifikan. Pengelompokan data bertujuan membagi data dalam beberapa kelompok (cluster) sedemikian hingga data-data dalam cluster yang sama (intra-cluster) memiliki derajat kesamaan yang tinggi, sementara data-data dalam cluster yang berbeda (inter-cluster) memiliki derajat kesamaan yang rendah. Model clustering data gangguan yang dirancang dengan metode algoritma k-means, model aplikasi ini dapat menampilkan gambaran dan menunjukkan pola sebaran data keluhan pelanggan.

Kata Kunci: Clustering, Data Mining, Algoritma K-Means, Matlab R2010a

Abstract

Data mining is the process of converting large amounts of data to generate connectivity between the contents of the data, the process of grouping the data (data clustering) becomes very important because of the increasing amount of data in text format that is quite significant. Grouping the data aims to divide the data into groups (clusters) such that the data in the same cluster (intra-cluster) has a high degree of similarity, while the data in different clusters (inter-cluster) have a low degree of similarity. Interruption of data clustering models designed by the method of k-means algorithm, the model of this application can display a picture and show the distribution pattern of customer complaint data.

Keywords: Clustering, Data Mining, K-Means algorithm, Matlab R2010a

A. PENDAHULUAN

A.1. Latar Belakang

Kemajuan teknologi saat ini membuat manusia hidup dalam suasana yang nyaman dan praktis. Hal ini dimungkinkan dengan adanya energi listrik. Dengan energi listrik, berbagai jenis peralatan listrik dapat diubah menjadi energi putar, panas, cahaya, serta sinyal audio-video, sesuai kebutuhan. Proses perubahan energi hingga listrik siap pakai di rumah-rumah atau di kantor-kantor membutuhkan biaya. Besarnya biaya yang harus disediakan tergantung dari jumlah tenaga listrik yang dimanfaatkan atau sering disebut dengan jumlah kWh terpakai.

TUL (Tata Usaha Langganan) adalah sistem pelayanan pelanggan yang meliputi kegiatan pelayanan kepada pelanggan atau calon pelanggan dan masyarakat lainnya yang membutuhkan tenaga listrik.

PLN sebagai sebuah penyedia jasa pelayanan umum yang mengawasi sektor ketenagalistrikan di Indonesia berkewajiban untuk selalu meningkatkan bentuk pelayanan kepada konsumen. Tidak terkecuali pada harapan konsumen di PT PLN (Persero) Area Pelayanan Cengkareng. Proses pengelompokan data (*data clustering*) menjadi sangat penting,

dikarenakan peningkatan jumlah data dalam format teks yang cukup signifikan belakangan ini. Pengelompokan data bertujuan membagi data dalam beberapa kelompok (*cluster*) sedemikian hingga data-data dalam *cluster* yang sama (*intra-cluster*) memiliki derajat kesamaan yang tinggi, sementara data-data dalam cluster yang berbeda (*inter-cluster*) memiliki derajat kesamaan yang rendah. Tidak ubahnya dengan PLN Area Pelayanan Cengkareng yang sangat membutuhkan aplikasi yang mampu mengelompokkan data keluhan pelanggan.

K-Means merupakan salah satu metode data clustering non hirarki yang berusaha mempartisi data yang ada ke dalam bentuk satu atau lebih *cluster*/kelompok. Metode ini mempartisi data ke dalam *cluster*/kelompok sehingga data yang memiliki karakteristik sama dikelompokkan ke dalam satu cluster yang sama.

Mengakomodasi pelanggan sehingga pelanggan mudah untuk menyampaikan keluhan-nya. Untuk itu diinginkan data keluhan pelanggan tersebut dikelompokkan ke dalam satu *cluster* yang sama. Dalam rangka menemukan pola – pola dan informasi dari sekumpulan koleksi data yang besar dapat dilakukan dengan menggunakan data mining. *Data mining* proses mengubah sejumlah besar data sampai

menghasilkan keterhubungan antar isi data serta dapat mendukung keputusan penting.

Dengan melihat permasalahan dan berdasarkan latar belakang masalah diatas, maka didapatkan suatu judul penelitian "MODEL CLUSTERING MENGGUNAKAN ALGORITMA K-MEANS PADA DATA KELUHAN PELANGGAN PT. PLN (STUDI KASUS : PT. PLN (PERSERO) DISTRIBUSI JAKARTA DAN TANGERANG)

A.2 Rumusan Masalah

1. Bagaimana melakukan pengolahan data keluhan pelanggan dengan menggunakan algoritma k-means.
2. Bagaimana membuat suatu aplikasi model *clustering* berdasarkan data keluhan pelanggan dengan menggunakan algoritma k-means

B. TINJAUAN PUSTAKA

B.1 Data Mining

Pengertian *Data Mining*

Secara sederhana data mining adalah penambangan atau penemuan informasi baru dengan mencari pola atau aturan tertentu dari sejumlah data yang sangat besar. Data mining juga disebut sebagai serangkaian proses untuk menggali nilai tambah berupa pengetahuan yang selama ini tidak diketahui secara manual dari suatu kumpulan data. Data mining, sering juga disebut sebagai *knowledge discovery in database* (KDD). KDD adalah kegiatan yang meliputi pengumpulan, pemakaian data, historis untuk menemukan keteraturan, pola atau hubungan dalam set data berukuran besar. Data mining adalah kegiatan menemukan pola yang menarik dari data dalam jumlah besar, data dapat disimpan dalam database, data *warehouse*, atau penyimpanan informasi lainnya

B.2 Clustering

Clustering termasuk metode yang sudah cukup dikenal dan banyak dipakai dalam data mining. Sampai sekarang para ilmuwan dalam bidang data mining masih melakukan berbagai usaha untuk melakukan perbaikan model *clustering* karena metode yang dikembangkan sekarang masih bersifat heuristik. Usaha-usaha untuk menghitung jumlah cluster yang optimal dan pengklasteran yang paling baik masih terus dilakukan

B.3 Algoritma K-Means

Algoritma K-Means merupakan satu algoritma yang mudah dan kerap digunakan di dalam teknik pengelompokan karena ia melibatkan pengiraan yang efisien dan tidak memerlukan banyak parameter. K-Means menggunakan k kelompok yang telah ditetapkan (k kelompok pertama sebagai centroid) dan secara berterusan akan melalui proses pengiraan titik tengah (min) sehingga sesuatu fungsi kriteria dicapai (kelompok adalah tetap). Di dalam teknik pengelompokan, pengiraan untuk membedakan di antara kelompok dilakukan menggunakan satu algoritma yang dipanggil fungsi yaitu tahap persamaan atau perbedaan

C. TUJUAN DAN MANFAAT PENELITIAN

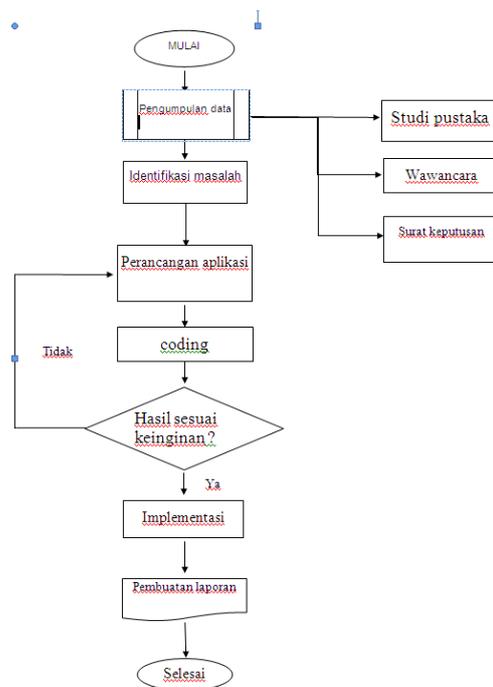
C.1. Tujuan Penelitian

untuk mengetahui pada tanggal berapa saja terjadi penumpukan keluhan dengan jenis gangguan yang berbeda dengan menerapkan clustering data menggunakan algoritma k-means.

C.2. Manfaat Penelitian

Memberikan informasi berupa data yang real mengenai keluhan pelanggan berdasarkan waktu terjadinya keluhan agar mempermudah dalam mengetahui pada tanggal-tanggal berapa sering terjadi pengeluhan oleh pelanggan

D. METODOLOGI PENELITIAN

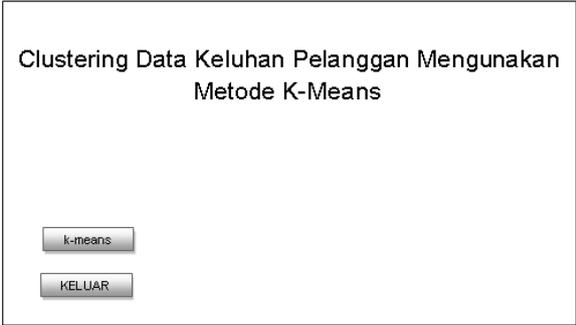


Gambar 1. Metodologi Penelitian

E. HASIL YANG TELAH DICAPAI

E.1. Perancangan Aplikasi Rancangan Sistem

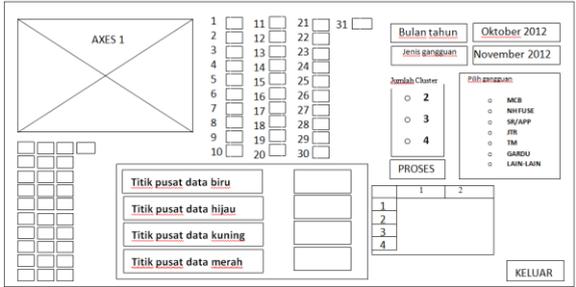
Rancangan yang dibuat merupakan rancangan sistem menggunakan metode K-Means dengan jumlah cluster 2, 3, dan 4. Hasil dari rancangan sistem ini berupa grafik dan pengelompokan tanggal dimana terjadinya penumpukan keluhan pelanggan yang terbentuk sesuai dengan jumlah cluster masing-masing. Langkah awal yang dilakukan untuk menjalankan program ini adalah dengan "Run" program. Kemudian tampilan perancangan awal dari program ini akan muncul seperti pada gambar dibawah ;



Gambar 2. Perancangan Tampilan awal program

Setelah itu kita akan memilih metode k-means untuk mengolah data. Apabila kita sudah memilih clustering dengan metode K-Means maka jendela clustering dengan metode K-Means akan muncul. Apabila kita memilih keluar maka kita akan meninggalkan program ini.

Kemudian langkah selanjutnya adalah memilih salah satu bulan yang ingin dieksekusi, apakah data bulan januari, february, maret, april, mei, juni, juli, agustus, september, oktober, november, desember pada tahun 2012. Kemudian memilih jenis gangguan apa yang ingin dianalisis, sehingga data gangguan pada bulan tersebut akan muncul. Jumlah keluhan akan terisi pada kotak-kotak sebanyak 31 sesuai jumlah hari pada satu bulannya.



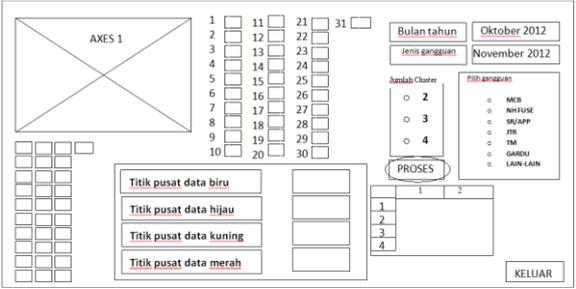
Gambar 3. Rancangan Tampilan Menu Utama

Data-data tiap jenis gangguan diasumsikan masing-masing mempunyai jumlah hari sebagai berikut:

Tabel 1. Tabel jumlah bulan dan hari dalam satu tahun

Bulan	Jumlah Hari
Januari	31 hari
Februari	28 hari
Maret	31 hari
April	30 hari
Mei	31 hari
Juni	30 hari
Juli	31 hari
Agustus	31 hari
September	30 hari
Oktober	31 hari
November	30 hari
Desember	31 hari

Langkah selanjutnya dalam proses clustering pada sistem aplikasi ini adalah menentukan jumlah cluster yang akan menjadi jumlah kelompok gangguan. Baru kemudian data bisa dieksekusi dengan mengklik tombol "Proses".



Gambar 4. Tombol eksekusi proses

Setelah mengeksekusi data tersebut maka akan muncul pengelompokan berdasarkan jumlah data gangguan perbulannya untuk masing-masing jenis gangguan.

Sebagaimana yang telah dijelaskan di atas masing-masing kategori clustering ini dibedakan menurut warna titik pusat klasternya, dimana setiap titik pusat mewakili sebuah kategori. Akan tetapi warna tersebut tidak konsisten mewakili sebuah kategori, kategori titik pusat warna dapat berubah pada proses eksekusi yang lain.

Range kelompok gangguan tinggi, sedang, cukup, dan rendah tidak di tentukan dari angka berapa hingga angka berapa. Akan tetapi range ini ditentukan berdasarkan hasil dari clustering dengan aturan-aturan dari metode yang digunakan.

Setiap jumlah cluster dari metode memiliki kategori yang berbeda. Masing-masing kategori clustering ini dibedakan menurut warna titik pusat klasternya, dimana setiap titik pusat mewakili sebuah kategori. Warna-warna pada titik pusat ini juga tidak konsisten mewakili sebuah kategori pada suatu

eksekusi. Pada eksekusi pengolahan data selanjutnya warna tiap titik pusat berkemungkinan mempunyai kategori yang berbeda

Titik Pusat Data Biru	<input type="text"/>
Titik pusat data Hijau	<input type="text"/>
Titik Pusat Data Kuning	<input type="text"/>
Titik Pusat Data Merah	<input type="text"/>

Gambar 5. perancangan keterangan titik pusat data

E.2. Hasil Uji Coba Sistem

Uji coba dilakukan dengan menggunakan data-data gangguan keluhan pelanggan PT-PLN (Persero) Distribusi Jakarta Dan Tangerang selama satu tahun. Kemudian data-data tersebut dikirimkan ke MATLAB dengan menggunakan Excel dan diberi nama variabel agar dapat diidentifikasi di dalam program matlab itu sendiri. Nama-nama dari variabel tersebut nantinya akan menjadi nama data matlab dengan ekstensi file *.mat.

Pada eksekusi data bulan desember dengan jenis gangguan MCB, jumlah cluster 4 dengan metode clustering K-Means didapat hasil sebagai berikut gambar 4.45 Hasil eksekusi data gangguan dengan jenis gangguan MCB jumlah cluster 4 metode K-Means (Lampiran).

Penjelasan warna titik pusat pada gambar 4.45

Tabel 2 Urutan kategori warna titik pusat

Warna Titik Pusat	Perbandingan	Urutan Katagori Keluhan
Titik pusat biru	>titik pusat hijau	Tinggi
Titik pusat hijau	>titik pusat kuning	Rendah
Titik pusat kuning	<titik pusat merah	Sangat Rendah
Titik pusat merah	<titik pusat biru	Sedang

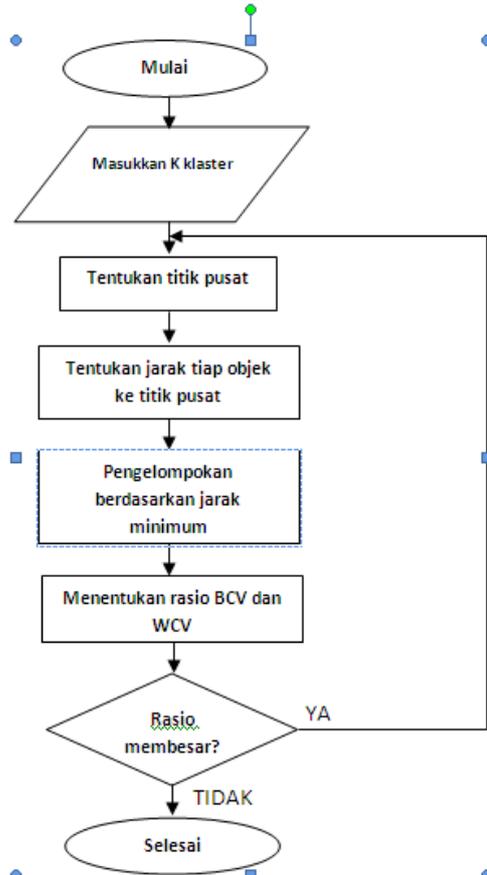
Berdasarkan tabel di atas titik pusat dengan kelompok banyaknya keluhan tertinggi berada pada titik pusat berwarna biru, titik pusat berwarna hijau berada pada kategori rendah, titik pusat berwarna kuning berada pada kategori sangat rendah sedangkan titik pusat berwarna merah berada pada kategori sedang.

Tampak bahwa jumlah gangguan dengan kriteria keluhan pelanggan tinggi yang diolah dengan metode K-Means dengan jumlah cluster 4 memiliki jumlah lebih sedikit dari pada jumlah gangguan dengan kriteria keluhan rendah dan sedang adalah jumlah yang paling banyak dibandingkan dengan kriteria lain. Akan

tetapi jumlah gangguan dengan kriteria keluhan cukup bertambah.

E.3. Perhitungan Manual Metode K-Means

Contoh Penghitungan Manual dengan Metode K-Means



Gambar 6. Flowchart metode K-Means

Contoh pengelompokan data keluhan pelanggan jenis gangguan MCB pada bulan November 2012 dengan metode K-Means:

Data keluhan pelanggan dengan jenis gangguan mcb disajikan dalam tabel berikut:

Tabel 3. Data Keluhan Pelanggan Jenis Gangguan MCB (nilai acak)

Tanggal	Mcb
1	4
2	0
3	1
4	1
5	5
6	1
7	5
8	4
9	0
10	3
11	1
12	5
13	0
14	1
15	2
16	1
17	2
18	1
19	3
20	1

Tanggal	Mcb
21	2
22	4
23	3
24	0
25	3
26	2
27	4
28	1
29	5
30	4
31	0

Langkah pertama untuk melakukan *clustering* K-Means adalah menentukan jumlah cluster/ kelompok dari data. Dalam kasus ini penulis akan membaginya menjadi 3 kelompok. Maka k=3.

Langkah kedua, dari masing-masing kelompok kita tentukan objek mana yang akan dijadikan titik pusat masing-masing kelompok. Misalkan kita tentukan:

- Titik pusat kelompok 1 = (1 , 4)
- Titik pusat kelompok 2 = (10 , 3)
- Titik pusat kelompok 3 = (21 , 2)

Langkah ketiga, kita tentukan pusat kelompok terdekat dari setiap objek. Objek tersebut akan ditetapkan sebagai anggota kelompok yang jaraknya terdekat langkah kedua.

Tentukan jarak terdekat ke titik pusat dengan menggunakan rumus :

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

Tabel 4. Penentuan kelompok metode K-Means iterasi pertama

Tanggal	Jarak ke pusat kelompok 1 (d1)	Jarak ke pusat kelompok 2 (d2)	Jarak ke pusat kelompok 3 (d3)	Jarak terdekat ke kelompok
1	0	9.055	20.099	1
2	4.123	8.544	19.104	1
3	3.605	7.280	18.027	1
4	4.242	6.324	17.029	1
5	4.123	5.385	16.278	1
6	5.830	4.472	15.033	2
7	6.082	3.605	14.317	2
8	7	2.236	13.152	2
9	8.944	3.162	12.165	2
10	9.055	0	11.045	2
11	10.440	2.236	10.049	2
12	11.045	2.828	9.486	2
13	12.649	4.242	8.246	2
14	13.341	4.472	7.071	2
15	14.142	5.099	6	2
16	15.297	6.324	5.099	3
17	16.124	7.071	4	3
18	17.262	8.246	3.162	3
19	18.027	9	2.236	3
20	19.235	10.198	1.414	3
21	20.099	11.045	0	3
22	21	12.041	2.236	3
23	22.022	13	2.828	3
24	23.345	14.317	3.605	3
25	24.020	15	4.123	3
26	25.079	16.031	5	3
27	26	17.029	6.324	3
28	27.166	18.110	7.071	3
29	28.017	19.104	8.544	3
30	29	20.024	9.219	3
31	30.265	21.213	10.198	3

Setelah melalui langkah diatas, langkah selanjutnya adalah menentukan rasio besaran antara *Between Cluster Variation* (BCV) dengan *Within Cluster Variation* (WCV), dimana jarak yang digunakan dalam menentukan BCV adalah jarak *Euclidean* dengan rumus berikut:

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

- Kelompok 1 (atau d1) = {1,2,3,4,5}
 - Kelompok 2 (atau d2) = {6,7,8,9,10,11,12,13,14,15}
 - Kelompok 3 (atau d3) = {16,17,18,19,20,21,22,23,24,26,27,28,29,30,31}
 - Rasio antar besaran Bertween Cluster Variation (BCV) dengan Within cluster variation (WCV)
 - Hitung nilai Between Cluster Variation (BCV)
 - BCV = kelompok1 ke kelompok2 + kelompok1 ke kelompok3 + kelompok2 ke kelompok3
- BCV = d(m1,m2)+d(m1,m3)+d(m2,m3)
= (9.055) + (20.099) + (11.045) = 40.399
 - WCV = (0)2 + (4.123) 2 + (3.605)2 + (4.242)2 + (4.123) 2 + (4.472)2 + (3.605)2 + (2.236)2 +

$$(3.162)^2 + (0)^2 + (2.236)^2 + (2.828)^2 + (4.242)^2 + (4.472)^2 + (5.099)^2 + (5.099)^2 + (4)^2 + (3.162)^2 + (2.236)^2 + (1.414)^2 + (2.236)^2 + (0)^2 + (2.236)^2 + (2.828)^2 + (3.605)^2 + (4.123)^2 + (5)^2 + (6.324)^2 + (7.071)^2 + (8.544)^2 + (9.219)^2 + (10.198)^2 = 668.892714$$

3. Maka rasionya adalah $BCV/WCV = 40.399/668.892714 = 0.060396$

Sampai pada langkah ini berarti kita telah melewati iterasi pertama, untuk memulai iterasi kedua maka kita tentukan lagi pusat-pusat kelompok yang baru.

- $m_1 =$ rata-rata jarak terdekat anggota-anggota kelompok 1 dengan pusat kelompok 1 pada iterasi pertama = (3, 3.218)
- $m_2 =$ rata-rata jarak terdekat anggota-anggota kelompok 2 dengan pusat kelompok 2 pada iterasi pertama = (9, 3.235)
- $m_3 =$ rata-rata jarak terdekat anggota-anggota kelompok 3 dengan pusat kelompok 3 pada iterasi pertama = (26, 4.691)

Selanjutnya, kita akan kembali melakukan perhitungan jarak terdekat dengan pusat kelompok baru seperti pada langkah ketiga.

$$(5.099)^2 + (1.964)^2 + (2.691)^2 + (1.215)^2 + (4.198)^2 + (3.015)^2 + (4.059)^2 + (6.856) = 636.150128$$

3. Maka rasionya adalah $BCV/WCV = 0.07248398447$

Sampai pada langkah ini artinya kita telah melewati iterasi yang kedua, langkah selanjutnya adalah menentukan kembali titik-titik pusat kelompok yang baru.

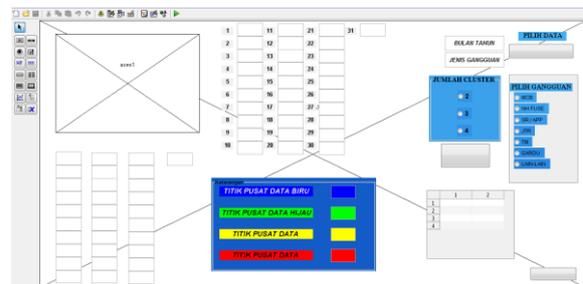
- $m_1 =$ rata-rata jarak terdekat anggota-anggota kelompok 1 dengan pusat kelompok 1 pada iterasi kedua = (5, 2.9901)
- $m_2 =$ rata-rata jarak terdekat anggota-anggota kelompok 2 dengan pusat kelompok 2 pada iterasi kedua = (13, 4.4187)
- $m_3 =$ rata-rata jarak terdekat anggota-anggota kelompok 3 dengan pusat kelompok 3 pada iterasi kedua = (24, 4.6665)

Baru kemudian kita kembali menentukan jarak terdekat data-data IPK dengan titik-titik pusat yang baru sampai rasio antara BCV dan WCV iterasi terakhir tidak lebih besar dari iterasi sebelumnya.

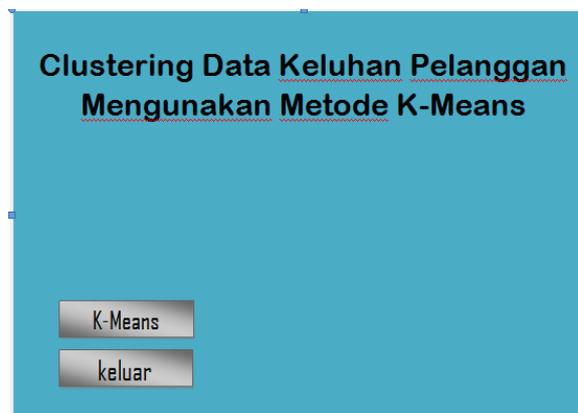
Tabel 5. Penentuan kelompok metode K-Means iterasi kedua

Objek	Jarak ke pusat kelompok 1	Jarak ke pusat kelompok 2	Jarak ke pusat kelompok 3	Jarak terdekat ke kelompok
1	2.147	8.036	25.009	1
2	3.369	7.711	24.454	1
3	2.218	6.402	23.294	1
4	2.433	5.476	22.307	1
5	2.678	4.372	21.002	1
6	3.730	3.741	20.337	1
7	4.378	2.667	19.002	1
8	5.060	1.259	18.013	2
9	6.808	3.235	17.635	2
10	7.003	1.027	16.089	2
11	8.301	2.999	15.447	2
12	9.174	3.480	14.003	2
13	10.505	5.144	13.820	2
14	11.221	5.476	12.554	2
15	12.061	6.125	11.324	2
16	13.187	7.348	10.659	2
17	14.052	8.094	9.393	2
18	15.163	9.273	8.810	3
19	16.001	10.002	7.201	3
20	17.144	11.224	7.044	3
21	18.041	12.063	5.678	3
22	19.016	13.022	4.059	3
23	20.001	14.001	3.443	3
24	21.245	15.344	5.099	3
25	22.001	16.001	1.964	3
26	23.032	17.044	2.691	3
27	24.012	18.016	1.215	3
28	25.098	19.131	4.198	3
29	26.060	20.077	3.015	3
30	27.011	21.013	4.059	3
31	28.184	22.236	6.856	3

HASIL TAMPILAN



Gambar 7. Menu Utama Sebelum Program DiRun



Gambar 8. Tampilan Awal Program

- $BCV = d(m_1, m_2) + d(m_1, m_3) + d(m_2, m_3) = (6.000) + (23.047) + (17.063) = 46.110696$
- $WCV = (2.147)^2 + (3.369)^2 + (2.218)^2 + (2.433)^2 + (2.678)^2 + (3.730)^2 + (4.378)^2 + (1.259)^2 + (3.235)^2 + (1.027)^2 + (2.999)^2 + (3.480)^2 + (5.144)^2 + (5.476)^2 + (6.125)^2 + (7.348)^2 + (8.094)^2 + (8.810)^2 + (7.201)^2 + (7.044)^2 + (5.678)^2 + (4.059)^2 + (3.443)^2 +$



Gambar 9. Tampilan Menu Utama Setelah Program DiRun



Gambar 14. Tampilan Setelah Memilih Jenis Gangguan dengan 3 Cluster



Gambar 10. tampilan setelah Ambil Data Pada Bulan November



Gambar 15. Tampilan Titik Pusat 3 cluster



Gambar 11. Tampilan setelah memilih Jenis Gangguan



Gambar 16. Tampilan Setelah Memilih Jenis Gangguan dengan 4 Cluster



Gambar 12. Tampilan Setelah Memilih Jenis Gangguan dengan 2 Cluster



Gambar 17. tampilan Titik Pusat 4 cluster



Gambar 13. Tampilan Titik Pusat 2 cluster



Gambar 18. tampilan Titik Pusat 4 cluster

F. KESIMPULAN DAN SARAN

F.1. Kesimpulan

Dapat mengetahui pada waktu kapan dan jenis gangguan terjadinya penumpukan keluhan dari pelanggan.

F.2. Saran

1. Data yang digunakan baru sebatas periode satu tahun, diharapkan dapat dibangun aplikasi pengolahan data ini dengan jumlah maksimum data yang tidak terbatas.
2. Pengelompokan baru berdasarkan satu metode clustering, diharapkan dapat dikembangkan dengan metode-metode clustering lainnya.
3. Input data pada aplikasi ini masih bersifat statis, menggunakan excel yang diinput ke matlab, diharapkan dapat dikembangkan aplikasi yang dapat menginput data secara dinamis dari ekstensi file tertentu.

Daftar Pustaka

1. Pelayanan Pelanggan, khususnya pengaduan dan Informasi Pelanggan, PT PLN (Persero) Distribusi Jakarta dan Tangerang
2. Berson, Alex, dkk. 2000. Building Data Mining Applications for CRM. McGraw-Hill
3. McFadden, Fred R. dan Jeffrey A. Hoffer. 1991. Database Management. Benjamin/Cummings Pub. Co.
4. Sahid. 2006. Panduan Praktis Matlab. Yogyakarta: Andi.
5. Susanto, Sani dan Dedy Suryadi. 2010. Pengantar Data Mining. Yogyakarta : Andi
6. K-Means Matlab Octave .<http://www.madsyair.net/2012/01/08/k-means-matlaboctave-code/>. 20 Januari 2012
7. Mathworks. [www.Mathworks.com](http://www.mathworks.com). 30 Oktober 2011
8. Membuat nilai acak di microsoft excel. <http://hadi.web.id/2009/02/membuat-nilai-acak-di-microsoft-excel.html>. 21 November 2011