

Data Mining Pengelompokan Pemakaian Listrik Terboros Menggunakan Metode *Clustering* (Studi Kasus PLN Area Binjai)

Ediman Manik

* Corresponding author : ediman.kaputama@gmail.com

Program Studi: Komputerisasi Akuntansi STMIK Kaputama

Jl. Veteran No. 4A – 9A, Binjai, Sumatera Utara

Abstract-- Clustering is a method used to classify a set of objects into classes consisting of the same objects. In other words clustering is the process of grouping objects based on the similarity of characteristics among the objects. In this thesis will be discussed about the clustering method by using K-Means algorithm that is used to determine the value of the electricity usage group as well as grouping it based on the level of power usage and the hours of the flame so that it can be grouped and known which ones are lost. The K-Means algorithm is one of the clustering methods that researchers will use to solve the problem of grouping. Starting from the determination of the grouping boundaries to the determination of each cluster so that the groups formed the data of electricity consumption adjusted with the level of usage. The result of this thesis is the value of data grouped by level of electricity usage grouping. And with the support of GUI programming MATLAB expected this program can be used easily. With this application can be seen the results of 200 data for cluster 1 as much as 44 with the Group of use of electricity in the group at the address is Binjai Timur and for Power group 450 and Jam on > 1100 and Cluster 2 as many as 60 Users of the listrik peloros at the address group is Binjai Kota and for Power 450 and 901-1100 Clock 3 and 96 for cluster 3 groups with License usage groups in the Address group are North Binjai and for Daya Power 450 and Jam 700-900.

Keywords: Clustering, Electricity Usage

Abstrak-- Emas merupakan alternatif yang cenderung dipilih kebanyakan orang untuk berinvestasi karena beberapa alasan seperti alasan keamanan, menguntungkan, mudah dicairkan, resiko rendah, tidak memerlukan dana besar, mudah dipindahkan, kepemilikan dan pengelolaan sendiri. Sulitnya memperkirakan turun naiknya harga emas membuat banyak enggan untuk berinvestasi, khususnya investasi emas. Peramalan atau prediksi adalah sama seperti teka-teki yang dipegang oleh banyak orang oleh karena penasaran dengan masa depan Automatic clustering and fuzzy logic relationship (ACFLR) merupakan salah satu metode dengan konsep fuzzy logic yang digunakan untuk pemodelan data time series. ACFLR diterapkan untuk mengetahui bagaimana tingkat keakuratannya dalam meramal harga emas sehingga dapat digunakan oleh investor untuk membuat perencanaan dalam investasi, alat pendukung keputusan untuk melakukan investasi. Berdasarkan hasil pembahasan, metode ACFLR mampu meramalkan harga emas dengan baik, dimana besarnya MAPE hanya 5,3% saja. Sehingga dapat disimpulkan metode ini layak dijadikan sebagai pendukung keputusan dalam berinvestasi emas.

Kata kunci: peramalan, data runtun, harga emas, ACFLR

PENDAHULUAN

Energi listrik merupakan salah satu faktor pendukung penting bagi kehidupan manusia karena banyak sekali peralatan yang biasa kita gunakan menggunakan listrik sebagai sumber energinya. seperti televisi, setrika, mesin cuci, handphone dan masih banyak lagi lainnya yang memerlukan energi listrik dalam menjalani kehidupan sehari-hari sebagai penyedia energi yang mempermudah dan membantu kegiatan manusia. PLN adalah salah satu lembaga pemerintah yang menjadi wadah dalam menyediakan energi listrik pada masyarakat. Namun dalam penyediaan energi listrik pada setiap masyarakat banyak ditemui kecurangan-kecurangan atau penyimpangan pemakaian aliran listrik yang dilakukan oleh perorangan. Dalam pemakai listrik yang ditemukan oleh pihak PLN banyak yang melakukan pemakaian daya listrik melebihi batas atau bisa dikatakan pemakaian menjadi boros melebihi ketentuan yang ditetapkan per pemilik. Dan dalam hal ini, pemakaian listrik ini dapat dikelompokkan sesuai dengan jumlah pemakaian pemilik tersebut. Pengelompokan data pemakaian listrik dengan menggunakan metode *clustering* dapat dilakukan dari semua pengelompokan-pengelompokan yang sudah ditentukan yang hasilnya adalah mengumpulkan pemakaian listrik yang ada pada PLN Binjai. Berdasarkan hasil pengelompokan tersebut, penulis akan menggali informasi penting yang dapat digunakan untuk mengetahui variabel-variabel apa saja yang menjadi pengelompokan pemakaian listrik dengan pemakaian terboros untuk *pengclusteran* pada data.

Data Mining

Data mining merupakan proses penerapan data dengan maksud untuk mengungkap pola-pola tersembunyi, dengan arti lain datamining adalah proses untuk penggalian pola-pola dari data. Data mining menjadi alat yang semakin penting untuk mengubah data menjadi sebuah informasi.

Menurut Pramudiono (2006, h, 3), Mengatakan bahwa "Data mining adalah serangkaian proses untuk menggali nilai tambah dari suatu kumpulan data berupa pengetahuan yang selama ini tidak diketahui secara manual".^[1] 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 bukanlah suatu bidang yang sama sekali baru, Salah satu kesulitan untuk mendefinisikan data mining adalah kenyataan bahwa data mining mewarisi banyak aspek dan teknik dari bidang-bidang ilmu yang sudah mapan terlebih dahulu.

Clustering

Clustering merupakan metode penganalisaan data, yang sering dimasukkan sebagai salah satu metode data mining, yang tujuannya adalah untuk mengelompokkan data dengan karakteristik yang sama. Menurut Baskoro (2010 h, 21) menyatakan bahwa "*Clustering* atau *clusterisasi* adalah salah satu alat bantu pada data *mining* yang bertujuan mengelompokkan obyek-obyek ke dalam *cluster-cluster*".^[2] *Cluster* adalah sekelompok atau sekumpulan obyek-obyek data yang *similar* satu sama lain dalam cluster yang sama dan *dissimilar* terhadap obyek-obyek yang berbeda *cluster*. Obyek akan dikelompokkan ke dalam satu atau lebih *cluster* sehingga obyek-obyek yang berada dalam satu *cluster* akan mempunyai kesamaan yang tinggi antara satu dengan lainnya. Obyek-obyek dikelompokkan berdasarkan prinsip memaksimalkan kesamaan obyek pada *cluster* yang sama dan memaksimalkan ketidaksamaan pada *cluster* yang berbeda. Kesamaan obyek biasanya diperoleh dari nilai-nilai atribut yang menjelaskan obyek data, sedangkan obyek-obyek data biasanya direpresentasikan sebagai sebuah titik dalam ruang multidimensi.

Algoritma K-Means

Algoritma *K-Means* digunakan untuk memecahkan masalah dalam pengelompokan data. Algoritma *k-means* disusun dalam bentuk *cluster* untuk menghasilkan suatu pengetahuan informasi yang terdapat pada setiap kelompok data yang digali. Menurut widyawati (2010 h, 7) menyatakan bahwa "Algoritma *k-means* merupakan algoritma yang membutuhkan parameter input sebanyak *k* dan membagi sekumpulan *n* objek kedalam *k cluster* sehingga tingkat kemiripan antar anggota dalam satu *cluster* tinggi sedangkan tingkat kemiripan dengan anggota pada *cluster* lain sangat rendah". Kemiripan anggota terhadap *cluster* diukur dengan kedekatan objek terhadap nilai *mean* pada *cluster*

atau dapat disebut sebagai *centroid cluster* atau pusat massa^{[3][4]}

MATLAB (*Matrix Laboratory*)

MATLAB adalah bahasa tingkat tinggi dan interaktif yang memungkinkan untuk melakukan komputasi secara intensif. Matlab merupakan suatu program komputer yang bisa membantu memecahkan berbagai masalah matematis yang kerap kita temui dalam bidang teknis. Menurut Agus Naba (2009, h, 39) menyatakan bahwa “MATLAB (*Matrix Laboratory*) adalah bahasa pemrograman tingkat tinggi dimana arti perintah dan fungsi-fungsinya bisa dimengerti dengan mudah, meskipun bagi seorang pemula”. Hal itu karena di dalam MATLAB, masalah dan solusi bisa diekspresikan dalam notasi-notasi matematis yang biasa dipakai. Pada awalnya MATLAB dimaksudkan sesuai dengan namanya, yaitu untuk menangani berbagai operasi matriks dan vektor menggunakan rutin-rutin dan library LINKPACK EISPACK dan BLAS, yang lebih efisien dalam menangani operasi matriks dan vektor.^{[5][6][7]}

METODE PENELITIAN

Tahapan Penelitian

- a. Definisi
Merupakan tahapan pendefinisian pokok permasalahan yang dimulai dari pencarian seputar Algoritma *K-Means* pada *data mining*.
- b. Menetapkan Tujuan
Yaitu memahami masalah dan menganalisis permasalahan dengan tujuan memperjelas dan menentukan *scope* permasalahan sehingga tercapailah tujuan.
- c. Mengumpulkan Data
Pengumpulan data dilakukan untuk memperoleh informasi yang dibutuhkan dalam rangka mencapai tujuan penelitian.
- d. Mengumpulkan dan Mempelajari Literatur
Metodologi penelitian yang digunakan adalah penelitian kepustakaan (*Library Research*) yaitu teknik pengumpulan data dengan mengadakan studi penelaahan terhadap buku-buku, literatur, catatan-catatan, artikel-artikel, laporan-laporan dan karya-karya ilmiah lain yang berhubungan dengan masalah yang akan dipecahkan yang menyangkut sistem kerja/operasional perusahaan atau instansi dimana penelitian dilaksanakan sesuai dengan bidang yang dipilih.
- e. Penentuan Metode penelitian
Agar diperoleh hasil yang baik maka penulis menerapkan metode penelitian asosiasi/hubungan yaitu suatu penelitian yang

bertujuan mengetahui hubungan antara dua variabel atau lebih.

- f. Analisis *Data mining* (Algoritma *K-Means*)
Setelah data terkumpul dan dipelajari langkah berikutnya yaitu analisa *Clustering* menggunakan Algoritma *K-Means* dengan variabel yang telah ditentukan yaitu Alamat, Daya dan Jam Nyala.
- g. Penentuan Hasil pengelompokkan
Setelah mengetahui hasil analisis dari *Clustering* (Algoritma *K-Means*) maka dapat ditentukan *Centroid* yang telah diperoleh.
- h. Test sistem menggunakan MatLab (*Matrix Laboratory*)
Yang merupakan tahapan terakhir dengan menguji hasil dan analisis *clustering* (Algoritma *K-means*) yang diimport ke MatLab (*Matrix Laboratory*).

Metode Pengumpulan Data

Data yang dibutuhkan untuk diinput kedalam pemrograman *MatLab* adalah data pemakaian listrik yang diambil adalah Alamat, Daya Dan Jam Nyala.

Data inputan yang akan diuji coba terdiri dari tiga variabel yaitu Alamat, Daya Dan Jam Nyala. Kemudian data-data pemakaian listrik di input kedalam *Microsoft Excel* sebagai *database* dengan nama “transformasi” dimana data yang akan diproses sebanyak 200 data.

HASIL DAN PEMBAHASAN

Dalam pembahasan antarmuka ini akan dijelaskan mengenai hasil perancangan program yang menggunakan *GUIDE Matlab*. Bisa dilihat sebagai berikut:

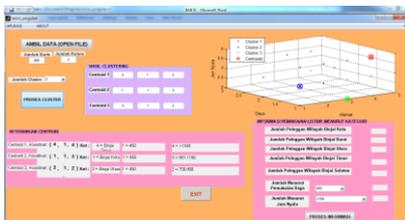
1. Menu *Home* atau *about* disini menampilkan *interface* awal yang berisi menu input berfungsi sebagai penginput data dan akan berhubungan dengan menu, menu cluster berfungsi sebagai proses pemilihan masing-masing *cluster* dan grafik. selain itu menu home



Gambar 1. Menu utama Dalam Clustering Pemakaian Listrik Terboros

2. Menu Input dalam *interface* ini berfungsi sebagai sarana penyimpanan data-data yang

berhubungan dalam metode *clustering* dengan algoritma *K-Means* atau disebut juga sebagai variabel, variable-variabel yang di inputkan adalah Alamat, Daya dan Jam Nyala. Dan ada inputan untuk group, inputan ini dilakukan berdasarkan hasil uji pada *MatLab*. Setelah semua data yang dibutuhkan diinput maka dapat meng-klik tombol proses yang berfungsi untuk menyimpan data ke dalam *tabel* dan berhubungan dengan menu tampil dapat dilihat pada gambar berikut:



Gambar 2. Menu Input Dalam Clustering Pemakaian Listrik

Keterangan output yaitu :

Centroid 4 1 4

Dapat diketahui bahwasanya pada *cluster1*. Kelompok pemakaian listrik terboros pada kelompok Alamat adalah Binjai Timur dan untuk kelompok Daya 450 dan Jam nyala >1100.

Centroid 1 1 3

Dapat diketahui bahwasanya pada *cluster2*. Kelompok pemakaian listrik terboros pada kelompok Alamat adalah Binjai Kota dan untuk kelompok Daya 450 dan Jam nyala 901-1100.

Centroid 3 1 2

Dapat diketahui bahwasanya pada *cluster3*. Kelompok pemakaian listrik terboros pada kelompok Alamat adalah Binjai Utara dan untuk kelompok Daya 450 dan Jam nyala 700-900.

KESIMPULAN

Diketahui hasil 200 data untuk cluster 1 sebanyak 44 dengan Kelompok pemakaian listrik terboros pada kelompok Alamat adalah Binjai Timur dan untuk kelompok Daya 450 dan Jam nyala >1100 dan Cluster 2 sebanyak 60 dengan Kelompok pemakaian listrik terboros pada kelompok Alamat adalah Binjai Kota dan untuk kelompok Daya 450 dan Jam nyala 901-1100 dan cluster 3 sebanyak 96 dengan Kelompok pemakaian listrik terboros pada kelompok Alamat adalah Binjai Utara dan untuk kelompok Daya 450 dan Jam nyala 700-900. Dapat diketahui hubungan antara variable Alamat, Daya dan Jam Nyala yaitu pemakaian listrik terboros terdapat pada cluster 1 yaitu Alamat adalah Binjai Timur dan untuk kelompok Daya 450 dan Jam nyala >1100. Pada metode *K-Means Clustering*, proses pengelompokan data dilakukan melalui sebuah

proses perulangan dimana perulangan akan dihentikan pada saat posisi cluster awal sudah sama dengan posisi *cluster* pada saat perulangan terakhir dilakukan.

DAFTAR PUSTAKA

- [1]Pramudiono (2013) Algoritma Pembelajaran Data Mining , CV. Andi Offset Yogyakarta..
- [2]Kusrini dan Luthfi Emha Taufiq, Algoritma Data Mining, Yogyakarta: Andi,2009.
- [3]Widyawati, dan iman paryudi., 2006. *Basis data*. Edisi Pertama, Penerbit Andi, Yogyakarta.
- [4]Nababan, D. (2017). SISTEM PENGONTROLAN PERSEDIAAN BARANG DENGAN METODE ECONOMIC ORDER QUANTITY (EOQ) MENGGUNAKAN ALGORITMA GENETIKA (Studi Kasus Gundaling Farm). *Journal Information System Development (ISD)*, 2(1).
- [5]Agus Naba, Data Mining : Konsep Dan Aplikasi Menggunakan Matlab, Penerbit C.V Andi, Yogyakarta, 2012
- [6]Prof. Dr. Jogiyanto HM,MBA,Akt., 2005, *Analisis dan Desain Sistem Informasi pendekatan terstruktur teori dan praktek aplikasi bisnis*, Edisi Ketiga, Penerbit Andi, Yogyakarta.
- [7]Tata Surbakti,S.Kom., MM., 2004. *Pemrograman Terstruktur*,Edisi Kedua, Andi , Yogyakarta.