PENERAPAN ANALYSIS CLUSTERING PADA PENJUALAN KOMPUTER DENGAN PERANCANGANAN APLIKASI DATA MINING MENGGUNAKAN ALGORITMA K-MEANS (STUDY KASUS TOKO TRI BUANA KOMPUTER KOTA SOLOK)

Surmayanti[1], Hari Marfalino[2], Ade Rahmi[3]
Fakultas Limu Komputer Universitas Putra Indonesia YPTK Padang

Email: Surmayanti@yahoo.co.id, hari.marfalino@yahoo.co.id, Dhecacute@ymail.com

Abstrak - Clustering merupakan salah satu metode Data Mining yang bersifat tanpa arahan (unsupervised), dan K-Means merupakan salah satu metode data Clustering non hirarki yang berusaha mempartisi data yang ada ke dalam bentuk satu atau lebih cluster/kelompok. Tujuan pembuatan aplikasi ini yaitumemberi informasi yang terkandung dari suatu data, sehingga usermendapatkan informasi yang belum didapatkandari data tersebut. Berdasarkan hasil uji coba,cepat atau lambatnya proses pengelompokan data (Clustering) dipengaruhi oleh spesifikasi suatu perangkat keras yang menjalankan aplikasi ini, operating sistem yang digunakan dan banyaknya data atau pembagiankelompok dari suatu data.

Kata kunci : Clustering, K-Means, Informasi

Clustering is one of data mining method without a trait (unsupervised), and K-Means is a kind of non hierarchical Clustering method that try to partition existing data into a kind or more cluster / group. The aim of this application is to provide information contained in a data, so user could get the information that hasn't been obtained from these data previously. Based on the test result, the speed of data clustrering depends on, a specification of a hardware device and operating system that performed the application and also the amount of data or the distribution of a data.

Key Words : Clustering, K-Means, Information

l. PENDAHULUAN

Latar Belakang

Dengan kemajuan teknologi informasi dewasa ini, kebutuhan akan informasi yang akurat sangat dibutuhkan dalam kehidupan sehari-hari, sehingga informasi akan menjadi suatu elemen penting dalam perkembangan masyarakat saat ini dan waktu mendatang. Namun kebutuhan informasi yang tinggi kadang tidak diimbangi dengan penyajian informasi yang memadai, sering kali informasi tersebut masih harus digali ulang dari data yang jumlahnya sangat besar.

Toko Tri Buwana Komputer merupakan toko yang bergerak dalam penjualan computer. Dalam rangka menghadapi persaingan bisnis terdapat beberapa permasalahan yang kerap muncul mengenai penjualan laptop. Toko Tri Buwana Computer sulit mendapatkan informasi-informasi strategis seperti tingkat penjualan per periode. Ketersediaan data penjualan yang besar DI Toko Tri Buwana Computer tidak digunakan

semaksimal mungkin, sehingga data penjualan tersebut tidak dimanfaatkan secara optimal.

Berdasarkan pada latar belakang diatas maka permasalahan diatas maka permasalahan yang ada yaitu belum adanya penerapan data mining untuk mengolah data penjualan computer. Berdasarkan kebutuhan diatas penulis mencoba memberi alternatif bantuan yang diwujudkan dalam sebuah penelitian dengan judul "**PENERAPAN ANALYSIS CLUSTERING PADA PENJUALAN KOMPUTER DENGAN PERANCANGAN APLIKASI DATA MINING MENGGUNAKAN ALGORITMA K-MEANS (STUDY KASUS TOKO TRI BUWANA KOMPUTER KOTA SOLOK)"**

Perumusan Masalah

Berdasarkan permasalahan diatas, maka rumusan masalah yang akan dibahas dalam penelitian ini adalah sebagai berikut :

1. Bagaimanakah aplikasi data mining mengefektifkan kinerja Toko Tri Buwana Komputer yang masih cenderung lambat?
2. Bagaimana aplikasi data mining memberikan arti yang sempurna dalam pengolahan data, karena dengan pengolahan data yang baik akan mudah mengakses data yang dibutuhkan?
3. Bagaimana aplikasi data mining ini dapat membantu pihak toko dalam memprediksi tingkat penjualan laptop di Toko Tri Buwana Komputer?

Hipotesa

Dari rumusan diatas maka dapat dikemukakan hipotesa yaitu, dengan menganalisa data penjualan laptop pada Toko Tri Buwana Komputer diharapkan membantu pihak toko dalam menganalisa dan mengambil keputusan serta dapat mengefektifkan kinerja toko yang masih cenderung lambat.

Tujuan Penelitian

Penelitian ini bertujuan untuk membangun sebuah perangkat lunak untuk memprediksi presentase suatu produk laptop, dimana perangkat lunak ini dapat mudah dioperasikan dan dipahami oleh pihak toko Tri Buwana Computer dalam menganalisa dan mengambil keputusan serta dapat mengoptimalkan dan mengefektifkan kinerja toko tersebut.

2. LANDASAN TEORI

Data mining adalah proses yang mempekerjakan satu atau lebih teknik pembelajaran komputer (machine learning) untuk menganalisis dan mengekstraksi pengetahuan (knowledge) secara otomatis (Fajar Astuti Hermawati, 2009).

Data mining adalah suatu istilah yang digunakan untuk menguraikan penemuan pengetahuan didalam *database*. *Data mining* adalah proses yang menggunakan teknik *statistic*, matematika, kecerdasan buatan, dan *machine learning* untuk mengekstraksi dan mengidentifikasi informasi yang bermanfaat dan pengetahuan yang terakit dari berbagai *database* besar (Turban,dkk. 2005) / (Kusrini & Luthfi,2009).

Pengelompokan Data Mining

Data Mining dibagi menjadi beberapa kelompok berdasarkan tugas yang dilakukan, yaitu (Larose, 2005) / (Kusrini & Luthfi,2009) :

1. Deskripsi

   Terkadang peneliti dan analisis secara sederhana ingin mencari cara untuk menggambarkan pola dan kecenderungan yang terdapat dalam data. Sebagai contoh, petugas pengumpulan suaramungkin tidak dapat menemukan keterangan atau fakta bahwa siapa yang tidak cukup professional akan sedikit didukung dalam pemilihan presiden. Deskripsi dari pola dan kecenderungan sering memberikan kemungkinan penjelasan untuk suatu pola atau kecenderungan.
2. Estimasi

   Estimasi hamper sama dengan klasifikasi, kecuali variable target estimasilebih kearah numeric dari pada ke arah kategori. Model dibangun menggunakan *record* lengkap yang menyediakan nilai dari variable target sebagai nilai prediksi. Selanjutnya , pada peninjauan berikutnya estimasi nilai dari variabel target dapat dibuat berdasarkan nilai variable prediksi. Sebagai contoh, akan dilakukan estimasi tekanan darah sistolik pada pasien rumah sakit berdasarkan umur pasien, jenis kelamin, berat badan, dan level sodium darah. Hubungan antara tekanan darah sistolik dan nilai variable prediksi dalam proses

pembelajararn akan menghasilkan model estimasi. Model estimasi yang digunakan untuk kasus baru lainnya.

3. Prediksi
   Prediksi hamper sama dengan klasifikasi dan estimasi, kecuali bahwa dalam prediksi nilai dari hasil akan ada dimasa mendatang.
   Contoh prediksi dalam bisnis dan penelitian adalah:
   a. Prediksi harga beras dalam tiga bulan yang akan datang.
   b. Prediksi presentase kenaikan kecelakaan lalu lintas tahun depan jika batas bawah kecepatan dinaikkan

   Beberapa metode dan teknik yang digunakan dalam klasifikasi dan estimasi dapat pula digunakan (untuk keadaan yang tepat) untuk prediksi.

4. Klasifikasi
   Dalam klasifikasi, terdapat target variable kategori. Sebagai contoh, penggolongan pendapatan dapat dipisahkan dalam tiga kategori, yaitu pendapatan tinggi, pendapatan sedang, dan pendapatan rendah.
   Contoh lain klasifikasi dalam bisnis dan penelitian adalah:
   a. Menentukan apakah suatu transaksi kartu kredit merupakan transaksi yang curang atau bukan.
   b. Memperkirakan apakah suatu pengajuan hipotek oleh nasabah merupakan suatu kredit yang baik atau buruk.
   c. Mendiagnosa penyakit seorang pasien untuk mendapatkan termasuk kategori apa.
5. Pengklusteran
   Pengklusteran mrerupakan pengelompokan *record* pengamatan , atau memperhatikan atau membentuk kelas objek-objek yang memiliki kemiripan.
   Kluster adalah kumpulan *record* yang memiliki kemiripan suatu dengan yang lainnya dan memiliki ketidakmiripan dengan *record* dalam kluster lain.

   Pengklusteran berbeda dengan klasifikasi yaitu tidak adanya variable target dalam pengklusteran. Pengklusteran tidak mencoba untuk melakukan klasifikasi, mengestimasi,atau memprediksi nilai dari variable target. Akan tetapi, algoritma pengklusteran mencoba untuk melakukan pembagian terhadap keseluruhan data menjadi kelompok-kelompok yang memiliki kemiripan (homogen), yang mana kemiripan dengan *record* dalam kelompok lain akan bernilai minimal.
   Contoh pengklusteran dalam bisnis dan penelitian adalah:
   a. Mendapatkan kelompok-kelompok konsumen untuk target pemasaran dari suatu produkbagi perusahaan yang tidak memiliki dana pemasaran yang besar.
   b. Untuk tujuan auditakutasi, yaitu melakukan pemisahan terhadap prilaku financial dalam baik dan mencurigakan.
   c. Melakukan pengklusteran terhadap ekspresi dari gen, dalam jumlah besar.
6. Asosiasi
   Tugas asosisasi dalam data mining adalah menentukan atribut yang muncul dalam suatu waktu. Dalam dunia bisnis lebih umum disebut analisis keranjang belanja.
   Contoh asosiasi dalam penelitian adalah:
   a. Meneliti jumlah pelanggan dari perusahaan telekomunikasi seluler yang diharapkan untuk memberikan respon positif terhadap penawaran upgrade layanan yang diberikan.
   b. Menemukan barang dalam supermarket yang dibeli secara bersamaan dan barang yang tidak pernah dibeli secara bersamaan.

Tantangan Dalam Data Mining

Tantangan dalam data mining meliputi (Fajar Astuti Hermawati, 2009) :

1. Scalability, yaitu besarnya ukuran basis data yang digunakan.
2. Dimensionality, yaitu banyaknya jumlah atribut dalam data yang akan diproses.
3. Complex and Heterogeneous Data, yaitu data yang kompleks dan mempunyai variasi yang beragam.
4. Data Quality, kualitas data yang akan diproses seperti data yang bersih dari noise, missing value, dsb.

5. Data Ownership and Distribution, yaitu siapa yang memiliki data dan bagaimana distribusinya.
6. Privacy Preservation, yaitu menjaga kerahasiaan data yang banyak dierapkan pada data nasabah perbankan.
7. Streaming Data, yaitu aliran data itu sendiri.

2.3.5 Kemajuan Dalam Data Mining

Kemajuan luar biasa yang terus berlanjut dalam bidang *data mining* didorong oleh beberapa factor, antara lain (Larose, 2005) / (Kusrini & Luthfi,2009).

1. Pertumbuhan yang cepat dalm kumpulan data.
2. Penyimpanan data dalam *data warehouse*, sehingga seluruh perusahaan memiliki akses ke dalam database yang andal.
3. Adanya peningkatan akses data melalui navigasi web dan intranet.
4. Tekanan kompetensi bisnis untuk meningkatkan penguasaan pasar dalam globalisasi ekonomi.
5. Perkembangan teknologi perangkat lunak untuk data mining (ketersediaan teknologi).
6. Perkembangan yang hebat dalam kemampuan komputasi dan pengembangan kapasitas media penyimpanan.

Tahap-tahap *Data Mining*

Istilah *data mining* dan *knowledge discovery in database* (KDD) sering kali digunakan secara bergantian unutk menjelaskan proses penggalian informasi tersembunyi dalam suatu basis data yang besar. Sebenarnya kedua istilah tersebut memiliki konsep yang berbeda,tetapi berkaitan satu sama lain. Dan salah satu tahapan dalam keseluruhan proses KDD adalah *data mining.* Proses KDD secara garis besar dapat dijelaskan sebagai berikut (Fayyad, 1996) / ((Kusrini & Luthfi,2009).

1. *Data Selection*

Pemilihan (seleksi) data dari sekumpulan data operasional perlu dilakukan sebelum tahap penggalian informasi dalam KDD dimulai dari data hasil seleksi yang akan digunakan untuk proses *data mining* disimpan dalam suatu berkas, terpisah dari basis *data operasional*.

2. *Pre-processing / Cleaning*

Sebelum proses *data mining* dapat dilaksanakan, perlu dilakukan proses *cleaning* pada data yang menjadi fokus KDD. Proses *cleaning* mencakup antara lain membuang duplikasi data, memeriksa data yang konsisten, dan memperbaiki kesalahan pada data, seperti kesalahan cetak (tipografi). Juga dilakukan proses *enrichment*, yaitu proses "memperkaya" data yang sudah ada. Dengan data atau informasi lain yang relevan dan diperlukan untuk KDD, seperti data atau informasi *eksternal*.

3. *Transformation*

Coding adalah proses *transformasi* pada data yang telah dipilih sehingga data tersebut sesuai untuk proses *data mining*. Proses *coding* dalam KDD merupakan proses kreatif dan sangat tergantung kepada jenis atau pola informasi yang akan dicari dalam bisnis data.

4. *Data Mining*

Data Mining adalah proses mencari pola atau informasi menarik dalam data terpilih dengan menggunakan teknik atau metode tertentu. Teknik, metode atau algoritma dalam *data mining* sangat bervariasi. Pemilihan metode atau algoritma yang tepat sangat bergantung pada tujuan dan proses KDD secara keseluruhan.

5. *Interpretation/Evaluation*

Pola informasi yang dihasilkan dari proses *data mining* perlu ditampilkan dalam bentuk yang mudah dimengerti oleh pihak yang berkepentingan. Tahap ini merupakan bagian dari proses KDD yang disebut *interpretation*. Tahap ini mencakup pemeriksaan apakah pola atau informasi yang ditemukan bertentangan dengan fakta atau hipotesis yang ada sebelumnya.

Clustering

Analisa *cluster* yaitu menemukan kumpulan objek hingga objek-objek dalam satu kelompok sama (atau punya hubungan) dengan yang lain dan berbeda (atau tidak berhubungan) dengan objek-objek dalam kelompok lain (Fajar Astuti Hermawati,2013).

Clustering adalah metode penganalisaan data, yang sering dimasukkan sebagai salah satu metode *Data Mining*, yang tujuannya adalah untuk mengelompokkan data dengan karakteristik yang sama ke suatu 'wilayah' yang sama dan data dengan karakteristik yang berbeda ke 'wilayah' yang lain.

Ada beberapa pendekatan yang digunakan dalam mengembangkan metode clustering.

Dua pendekatan utama adalah clustering dengan pendekatan partisi dan clustering dengan pendekatan hirarki. Clustering dengan pendekatan partisi atau sering disebut dengan *partition-based clustering* mengelompokkan data dengan memilah-milah data yang dianalisa ke dalam cluster-cluster yang ada. Clustering dengan pendekatan hirarki atau sering disebut dengan *hierarchical clustering* mengelompokkan data dengan membuat suatu hirarki berupa dendogram dimana data yang mirip akan ditempatkan pada hirarki yang berdekatan dan yang tidak pada hirarki yang berjauhan.

Tujuan Clustering

Ada pun tujuan dari data clustering ini adalah untuk meminimalisasikan *objective function* yang diset dalam proses clustering, yang pada umumnya berusaha meminimalisasikan variasi di dalam suatu cluster dan memaksimalisasikan variasi antar cluster (Jurnal Sistem dan Informatika Vol. 3, 2007).

K-means
Defenisi K-means

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 yang sama dikelompokkan ke dalam satu cluster yang sama dan data yang mempunyai karakteristik yang berbeda dikelompokkan ke dalam kelompok yang lain. (Jurnal Sistem dan Informatika Vol. 3, 2007).

K-means *Clustering*

Menggunakan pendekatan *partitional clustering* . Tiap *cluster* dihubungkan dengan sebuah *centroid* (titik pusat). Tiap titik pusat ditempatkan ke dalam *cluster* dengan *centroid* terdekat. Jumlah *cluster*, K, harus ditentukan. Algoritma dasarnya sangat sederhana, yaitu (Fajar Astuti Hermawati, 2013) :

1. Pilih K titik sebagai *centroid* awal
2. Ulangi
3. Bentuk K *cluster* dengan menempatkan semua titik yang terdekat.
4. Ulangi perhitungan *centroid* dari tiap *cluster*.
5. Sampai *centroid* tidak berubah.

3**. Analisa dan Hasil**

1. Analisa Sistem

Untuk analisa sistem, penulis menggunakan aturan asosiasi dimana bentuk umum dari aturan asosiasi adalah :

Algoritma pengelompokkan data :

a. Ambil nilai jarak tiap pusat cluster dengan data
b. Cari nilai jarak terkecil
c. Kelompokkan data dengan pusat cluster yang memiliki jarak terkecil.

1. Penentuan pusat cluster baru

Untuk mendapatkan pusat cluster baru bisa dihitung dari rata-rata nilai anggota cluster dan pusat cluster. Pusat cluster yang baru digunakan untuk melakukan iterasi selanjutnya, jika hasil yang didapatkan belum konvergen. Proses iterasi akan berhenti jika telah memenuhi maksimum iterasi yang dimasukkan oleh user atau hasil yang dicapai sudah konvergen (pusat cluster baru sama dengan pusat cluster lama).

Algoritma penentuan pusat cluster :

a. Cari jumlah anggota tiap cluster
b. Hitung pusat baru dengan rumus

$$\text{Pusat cluster baru} = \frac{x1 + x2 + x3 \ldots . + xn + xp}{\text{jumlah} + 1}$$

Dimana :
X1, X2, X3,.... Xn = anggota cluster
Xp = pusat lama

Dari aturan-aturan asosiasi tersebut, penulis merancang sistem dengan melakukan perhitungan terhadap data yang telah didapat. Data yang masuk dalam perhitungan, penulis menggunakan beberapa data transaksi saja sebagai sampel dari perancangan sistem.

Tabel Transaksi Penjualan

No	Nama Barang	Jumlah Barang	Jumlah Barang Terjual
1	NB Toshiba	34	17
2	Thosiba dual core	35	24
3	Thosiba core I3	55	31
4	NB Acer	27	24
5	Acer DC	36	24
6	Acer core I3	31	26
7	NB HP	14	5
8	HP DC	25	10
9	HP Core I3	13	7
10	NB Axioo	27	11
11	Axioo Qudd Core	23	11
12	Axioo core I5	19	11
13	NB Asus	14	2
14	Asus Dc	45	26
15	Dell Dc	23	13
16	Lenovo Core I3	16	7
17	NB Lenovo	20	5
18	Lenovo Dc	19	11

Pada tahap ini akan dilakukan proses utama yaitu segmentasi atau pengelompokkan data Penjualan Barang yang diakses dari database, yaitu sebuah metode clustering algoritma K-Means.
Dari banyak data penjualan yang diperoleh, Percobaan dilakukan dengan menggunakan parameter-parameter berikut :

Jumlah cluster : 2
Jumlah data : 18
Jumlah atribut : 2

Pada table dibawah ini merupakan sampel data yang digunakan untuk melakukan percobaan perhitungan manual.

2. Proses Iterasi ke-1

1. Penentuan pusat awal cluster

Pusat awal cluster atau centroid didapatkan secara random, untuk penentuan awal cluster di asumsikan :

Pusat Cluster 1: **(37 ,24)**
Pusat Cluster 2: **(19,8)**

2. Perhitungan jarak pusat cluster

Untuk mengukur jarak antara data dengan pusat cluster digunakan Euclidian distance, kemudian akan didapatkan matrik jarak sebagai berikut:

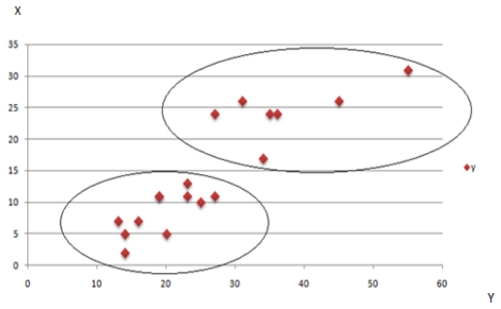
Rumus euclidian distance :

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

X = Pusat cluster
Y = data

Dari 10 data yang dijadikan sampel telah dipilih pusat awal cluster yaitu **C1 (37 , 24), dan C2 (19 , 8).** Lalu dilakukan penghitungan jarak dari sisa sampel data dengan pusat cluster yang dimisalkan dengan **M(a,b),** dimana **a** merupakan jumlah stock, dan **b** jumlah terjual yang diperkecil menjadi angka puluhan juta agar cara penghitungan lebih mudah.

1. M1 = (34,17)
2. M2 = (35,24)
3. M3 = (55,31)
4. M4 = (27, 24)
5. M5 = (36,24)
6. M6 = (31,26)
7. M7 = (14,5)
8. M8 = (25,10)
9. M9 = (13,7)
10. M10 = (27,11)
11. M11 = (23,11)
12. M12 = (19,11)
13. M13 = (14,2)
14. M14 = (45,26)
15. M15 = (23,13)
16. M16 = (16,7)
17. M17= (20,5)
18. M18 = (19,11)



Gambar Hasil Diagram Clustering

Hitung Euclidean distance dari semua data kesetiap titik pusat pertama :

Dengan cara yang sama hitung jarak titik ketitik pusat ke-2 dan kita akan mendapatkan :

$$D11=\sqrt{(M_{1x}-C_{1x})^2+(M_{1y}-C_{1y})^2}=\sqrt{(34-19)^2+(17-8)^2}=\sqrt{58}=7{,}61$$
$$D12=\sqrt{(M_{2x}-C_{1x})^2+(M_{2y}-C_{1y})^2}=\sqrt{(35-37)^2+(24-24)^2}=\sqrt{4}=2$$
$$D13=\sqrt{(M_{3x}-C_{1x})^2+(M_{3y}-C_{1y})^2}=\sqrt{(55-37)^2+(31-24)^2}=\sqrt{373}=19{,}31$$
$$D14=\sqrt{(M_{4x}-C_{1x})^2+(M_{4y}-C_{1y})^2}=\sqrt{(27-37)^2+(24-24)^2}=\sqrt{100}=10$$
$$D15=\sqrt{(M_{5x}-C_{1x})^2+(M_{5y}-C_{1y})^2}=\sqrt{(36-37)^2+(24-24)^2}=\sqrt{1}=1$$
$$D16=\sqrt{(M_{61x}-C_{1x})^2+(M_{6y}-C_{1y})^2}=\sqrt{(31-37)^2+(26-24)^2}=\sqrt{40}=6{,}32$$
$$D17=\sqrt{(M_{7x}-C_{1x})^2+(M_{7y}-C_{1y})^2}=\sqrt{(14-37)^2+(5-24)^2}=\sqrt{890}=29.83$$
$$D18=\sqrt{(M_{8x}-C_{1x})^2+(M_{8y}-C_{1y})^2}=\sqrt{(25-37)^2+(10-24)^2}=\sqrt{340}=18.43$$
$$D19=\sqrt{(M_{9x}-C_{1x})^2+(M_{9y}-C_{1y})^2}=\sqrt{(13-37)^2+(7-24)^2}=\sqrt{865}=29{,}41$$
$$D110=\sqrt{(M_{10x}-C_{1x})^2+(M_{10y}-C_{1y})^2}=\sqrt{(27-37)^2+(11-24)^2}=\sqrt{269}=16{,}40$$
$$D111=\sqrt{(M_{11x}-C_{1x})^2+(M_{11y}-C_{1y})^2}=\sqrt{(23-37)^2+(11-24)^2}=\sqrt{365}=19{,}10$$
$$D112=\sqrt{(M_{12x}-C_{1x})^2+(M_{12y}-C_{1y})^2}=\sqrt{(19-37)^2+(11-24)^2}=\sqrt{493}=22{,}20$$
$$D113=\sqrt{(M_{13x}-C_{1x})^2+(M_{13y}-C_{1y})^2}=\sqrt{(14-37)^2+(2-24)^2}=\sqrt{1013}=31.82$$
$$D114=\sqrt{(M_{14x}-C_{1x})^2+(M_{14y}-C_{1y})^2}=\sqrt{(45-37)^2+(26-24)^2}=\sqrt{68}=8.24$$
$$D115=\sqrt{(M_{15x}-C_{1x})^2+(M_{15y}-C_{1y})^2}=\sqrt{(23-37)^2+(13-24)^2}=\sqrt{317}=17{,}80$$
$$D116=\sqrt{(M_{16x}-C_{1x})^2+(M_{15y}-C_{1y})^2}=\sqrt{(16-37)^2+(7-24)^2}=\sqrt{730}=27{,}01$$
$$D117=\sqrt{(M_{17x}-C_{1x})^2+(M_{17y}-C_{1y})^2}=\sqrt{(20-37)^2+(5-24)^2}=\sqrt{650}=25{,}49$$
$$D118=\sqrt{(M_{18x}-C_{1x})^2+(M_{18y}-C_{1y})^2}=\sqrt{(19-37)^2+(11-24)^2}=\sqrt{493}=22{,}20$$

$$D21=\sqrt{(M_{1x}-C_{1x})^2+(M_{1y}-C_{1y})^2}=\sqrt{(34-19)^2+(17-8)^2}=\sqrt{306}=17{,}49$$
$$D22=\sqrt{(M_{2x}-C_{1x})^2+(M_{2y}-C_{1y})^2}=\sqrt{(35-19)^2+(24-8)^2}=\sqrt{512}=22{,}62$$
$$D23=\sqrt{(M_{3x}-C_{1x})^2+(M_{3y}-C_{1y})^2}=\sqrt{(55-19)^2+(31-8)^2}=\sqrt{1825}=42{,}72$$
$$D24=\sqrt{(M_{4x}-C_{1x})^2+(M_{4y}-C_{1y})^2}=\sqrt{(27-19)^2+(24-8)^2}=\sqrt{320}=17{,}88$$
$$D25=\sqrt{(M_{5x}-C_{1x})^2+(M_{5y}-C_{1y})^2}=\sqrt{(36-19)^2+(24-8)^2}=\sqrt{545}=23{,}34$$
$$D26=\sqrt{(M_{61x}-C_{1x})^2+(M_{6y}-C_{1y})^2}=\sqrt{(31-19)^2+(26-8)^2}=\sqrt{468}=21{,}63$$
$$D27=\sqrt{(M_{7x}-C_{1x})^2+(M_{7y}-C_{1y})^2}=\sqrt{(14-19)^2+(5-8)^2}=\sqrt{34}=5.83$$
$$D28=\sqrt{(M_{8x}-C_{1x})^2+(M_{8y}-C_{1y})^2}=\sqrt{(25-19)^2+(10-8)^2}=\sqrt{40}=6.32$$
$$D29=\sqrt{(M_{9x}-C_{1x})^2+(M_{9y}-C_{1y})^2}=\sqrt{(13-19)^2+(7-8)^2}=\sqrt{37}=6{,}08$$
$$D210=\sqrt{(M_{10x}-C_{1x})^2+(M_{10y}-C_{1y})^2}=\sqrt{(27-19)^2+(11-8)^2}=\sqrt{73}=8{,}54$$
$$D211=\sqrt{(M_{11x}-C_{1x})^2+(M_{11y}-C_{1y})^2}=\sqrt{(23-19)^2+(11-8)^2}=\sqrt{25}=5$$
$$D212=\sqrt{(M_{12x}-C_{1x})^2+(M_{12y}-C_{1y})^2}=\sqrt{(19-19)^2+(11-8)^2}=\sqrt{9}=3$$
$$D213=\sqrt{(M_{13x}-C_{1x})^2+(M_{13y}-C_{1y})^2}=\sqrt{(14-19)^2+(2-29)^2}=\sqrt{61}=7.81$$
$$D214=\sqrt{(M_{14x}-C_{1x})^2+(M_{14y}-C_{1y})^2}=\sqrt{(45-19)^2+(26-29)^2}=\sqrt{1000}=31.62$$
$$D215=\sqrt{(M_{15x}-C_{1x})^2+(M_{15y}-C_{1y})^2}=\sqrt{(23-19)^2+(13-29)^2}=\sqrt{41}=6{,}40$$
$$D216=\sqrt{(M_{16x}-C_{1x})^2+(M_{15y}-C_{1y})^2}=\sqrt{(16-19)^2+(7-29)^2}=\sqrt{10}=3{,}16$$
$$D217=\sqrt{(M_{17x}-C_{1x})^2+(M_{17y}-C_{1y})^2}=\sqrt{(20-19)^2+(5-29)^2}=\sqrt{10}=3{,}16$$
$$D218=\sqrt{(M_{18x}-C_{1x})^2+(M_{18y}-C_{1y})^2}=\sqrt{(19-19)^2+(11-29)^2}=\sqrt{9}=3$$

Dari hasil penghitungan Euclidean distance, kita dapat membandingkan :

Antara C1 dan C2, C1 menerangkan barang yang tidak laris sedangkan C2 menerangkan barang yang laris. Disini penulis hanya mengambil dua cluster karena penulis ingin membatasinya.

Tabel Hasil Iterasi 1

	C1	C2
M1	7,61	17,49
M2	2	22,62
M3	19,31	42,72
M4	10	17,88
M5	1	23,34
M6	6,32	21,63
M7	29,83	5,83
M8	18,43	6,32
M9	29,41	6,08
M10	16,40	8,54
M11	19,10	5
M12	22,20	3
M13	31,82	7,81
M14	8,24	31,62
M15	17,80	6.40
M16	27,01	3.16
M17	25,49	3,16
M18	22,20	3

{M1, M2,M3,M4,M5,M6,M14}: Anggota C1

{,M7,M8,M9,M10,M11,M12,M13,M15,M16 ,17,M18} : Anggota C2

Kesimpulan : M1,M2,M3,M4,M5,M6,M14 merupakan anggota dari cluster 1 dan M7,M8,M9,M10,M11,M12,M15,M16,M17,M 18 merupakan anggota dari cluster 2.

3. Proses Iterasi ke-2

1. Hitung titik pusat baru

Tentukan posisi centroid baru (Ck) dengan cara menghitung nilai rata-rata dari data-data yang ada pada centroid yang sama.

$$C_k=\left(\frac{1}{n_k}\right)\sum d_1$$

Dimana nk adalah jumlah dokumen dalam cluster k dan di adalah dokumen dalam cluster k.

$$C1 = \frac{34+35+55+27+36+31+45}{7},$$
$$\frac{17+24+31+24+24+26+26}{7}$$
$$= (263/7 =),(172/7 =)$$
$$= (37.5,24.5)$$

$$C2 = \frac{14+25+13+27+23+19+14+23+16+20+19}{11},$$
$$\frac{5+10+7+11+11+11+2+13+7+5+11}{11}$$
$$= (213/11 =),(93/11 =)$$
$$= (19.3,8.4)$$

Sehingga didapatkan titik pusat atau centroid yang baru yaitu :

C1= (37,5 , 24,5) C2= (19 ,3 , 8,4)

2. Perhitungan jarak pusat cluster

Hitung Euclidean distance dari semua data ketitik pusat yang baru (C1, C2) seperti yang telah dilakukan pada tahap 1. Setelah hasil perhitungan kita dapatkan, kemudian bandingkan hasil tersebut. Jika hasil posisi cluster pada iterasi ke 2 sama dengan posisi iterasi pertama, maka proses dihentikan, namun jika tidak proses dilanjutkan ke iterasi ke 3.

$D11 = \sqrt{(M_{1x}-C_{2x})^2+(M_{1y}-C_{2y})^2} = \sqrt{(34-37,5)^2+(17-24,5)^2} = \sqrt{68,5} = 8,27$
$D12 = \sqrt{(M_{2x}-C_{2x})^2+(M_{2y}-C_{2y})^2} = \sqrt{(35-37,5)^2+(24-24,5)^2} = \sqrt{6,5} = 2,54$
$D13 = \sqrt{(M_{3x}-C_{2x})^2+(M_{3y}-C_{2y})^2} = \sqrt{(55-37,5)^2+(31-24,5)^2} = \sqrt{348,5} = 18,66$
$D14 = \sqrt{(M_{4x}-C_{2x})^2+(M_{4y}-C_{2y})^2} = \sqrt{(27-37,5)^2+(24-24,5)^2} = \sqrt{110,5} = 10,51$
$D15 = \sqrt{(M_{5x}-C_{2x})^2+(M_{5y}-C_{2y})^2} = \sqrt{(36-37,5)^2+(24-24,5)^2} = \sqrt{2,5} = 1,58$
$D16 = \sqrt{(M_{6x}-C_{2x})^2+(M_{6y}-C_{2y})^2} = \sqrt{(31-37,5)^2+(26-24,5)^2} = \sqrt{44,5} = 6,67$
$D17 = \sqrt{(M_{7x}-C_{12x})^2+(M_{7y}-C_{2y})^2} = \sqrt{(14-37,5)^2+(5-24,5)^2} = \sqrt{932,5} = 30,53$
$D18 = \sqrt{(M_{8x}-C_{12})^2+(M_{8y}-C_{2y})^2} = \sqrt{(25-37,5)^2+(10-24,5)^2} = \sqrt{366,5} = 19.14$
$D19 = \sqrt{(M_{9x}-C_{2x})^2+(M_{9y}-C_{1y})^2} = \sqrt{(13-37,5)^2+(7-24,5)^2} = \sqrt{906,5} = 30,10$
$D110 = \sqrt{(M_{10x}-C_{2X})^2+(M_{10y}-C_{2y})^2} = \sqrt{(27-37,5)^2+(11-24,5)^2} = \sqrt{292,5} = 17,10$
$D111 = \sqrt{(M_{11x}-C_{2x})^2+(M_{11y}-C_{2y})^2} = \sqrt{(23-37,5)^2+(11-24,5)^2} = \sqrt{392,5} = 19,81$
$D112 = \sqrt{(M_{12x}-C_{2x})^2+(M_{12y}-C_{2y})^2} = \sqrt{(19-37,5)^2+(11-24,5)^2} = \sqrt{524,5} = 2,29$
$D113 = \sqrt{(M_{13x}-C_{2x})^2+(M_{13y}-C_{2y})^2} = \sqrt{(14-37,5)^2+(2-24,5)^2} = \sqrt{1058,5} = 32.53$
$D114 = \sqrt{(M_{14x}-C_{2x})^2+(M_{14y}-C_{2y})^2} = \sqrt{(45-37,5)^2+(26-24,5)^2} = \sqrt{58,5} = 7,64$
$D115 = \sqrt{(M_{15x}-C_{2x})^2+(M_{15y}-C_{2y})^2} = \sqrt{(23-37,5)^2+(13-24,5)^2} = \sqrt{342,5} = 18,50$
$D116 = \sqrt{(M_{16x}-C_{2x})^2+(M_{16y}-C_{2y})^2} = \sqrt{(16-37,5)^2+(7-24,5)^2} = \sqrt{768,5} = 27,72$
$D117 = \sqrt{(M_{17x}-C_{2x})^2+(M_{17y}-C_{2y})^2} = \sqrt{(20-37,5)^2+(5-24,5)^2} = \sqrt{686,5} = 26,20$
$D118 = \sqrt{(M_{18x}-C_{2x})^2+(M_{18y}-C_{2y})^2} = \sqrt{(19-37,5)^2+(11-24,5)^2} = \sqrt{524,5} = 22.90$

$D21 = \sqrt{(M_{1x}-C_{1x})^2+(M_{1y}-C_{1y})^2} = \sqrt{(34-19,3)^2+(17-8,4)^2} = \sqrt{290,05} = 17,03$
$D22 = \sqrt{(M_{2x}-C_{1x})^2+(M_{2y}-C_{1y})^2} = \sqrt{(35-19,3)^2+(24-8,4)^2} = \sqrt{489,85} = 22,13$
$D23 = \sqrt{(M_{3x}-C_{1x})^2+(M_{3y}-C_{1y})^2} = \sqrt{(55-19,3)^2+(31-8,4)^2} = \sqrt{1785,25} = 42,25$
$D24 = \sqrt{(M_{4x}-C_{1x})^2+(M_{4y}-C_{1y})^2} = \sqrt{(27-19,3)^2+(24-8,4)^2} = \sqrt{302,65} = 17,39$
$D25 = \sqrt{(M_{5x}-C_{1x})^2+(M_{5y}-C_{1y})^2} = \sqrt{(36-19,3)^2+(24-8,4)^2} = \sqrt{522,25} = 22,85$
$D26 = \sqrt{(M_{6x}-C_{1x})^2+(M_{6y}-C_{1y})^2} = \sqrt{(31-19,3)^2+(26-8,4)^2} = \sqrt{446,65} = 21,13$
$D27 = \sqrt{(M_{7x}-C_{1x})^2+(M_{7y}-C_{1y})^2} = \sqrt{(14-19,3)^2+(5-8,4)^2} = \sqrt{39,65} = 6,29$
$D28 = \sqrt{(M_{8x}-C_{1x})^2+(M_{8y}-C_{1y})^2} = \sqrt{(25-19,3)^2+(10-8,4)^2} = \sqrt{8,26} = 2,87$
$D29 = \sqrt{(M_{9x}-C_{1x})^2+(M_{9y}-C_{1y})^2} = \sqrt{(13-19,3)^2+(7-8,4)^2} = \sqrt{41,65} = 6,45$
$D210 = \sqrt{(M_{10x}-C_{1x})^2+(M_{10y}-C_{1y})^2} = \sqrt{(27-19,3)^2+(11-8,4)^2} = \sqrt{66,05} = 8,12$
$D211 = \sqrt{(M_{11x}-C_{1x})^2+(M_{11y}-C_{1y})^2} = \sqrt{(23-19,3)^2+(11-8,4)^2} = \sqrt{20,45} = 4,52$
$D212 = \sqrt{(M_{12x}-C_{1x})^2+(M_{12y}-C_{1y})^2} = \sqrt{(19-19,3)^2+(11-8,4)^2} = \sqrt{6,85} = 2,61$
$D213 = \sqrt{(M_{13x}-C_{1x})^2+(M_{13y}-C_{1y})^2} = \sqrt{(14-19,3)^2+(2-8,4)^2} = \sqrt{69,05} = 8,30$
$D214 = \sqrt{(M_{14x}-C_{1x})^2+(M_{14y}-C_{1y})^2} = \sqrt{(45-19,3)^2+(26-8,4)^2} = \sqrt{970,25} = 31,1$
$D215 = \sqrt{(M_{15x}-C_{1x})^2+(M_{15y}-C_{1y})^2} = \sqrt{(23-19,3)^2+(13-8,4)^2} = \sqrt{34,85} = 5,90$
$D216 = \sqrt{(M_{16x}-C_{1x})^2+(M_{16y}-C_{1y})^2} = \sqrt{(16-19,3)^2+8,4)^2} = \sqrt{12,85} = 3,58$
$D217 = \sqrt{(M_{17x}-C_{1x})^2+(M_{17y}-C_{1y})^2} = \sqrt{(20-19,3)^2+(5-8,4)^2} = \sqrt{12,05} = 3,47$
$D218 = \sqrt{(M_{18x}-C_{1x})^2+(M_{18y}-C_{1y})^2} = \sqrt{(19-19,3)^2+(11-8,4)^2} = \sqrt{6,85} = 2,61$

Dengan cara yang sama hitung jarak tiap titik ketitik pusat baru ke-2 dan kita akan mendapatkan :

Tabel Hasil Iterasi 2

	C1	C2
M1	13,27	17,03
M2	2,54	22,13
M3	18,66	42,25
M4	10,51	17,39
M5	1,58	22,85
M6	6,67	21,13
M7	30,53	6,29
M8	19,14	2,87
M9	30,10	6,45
M10	17,10	8,12
M11	19,81	4,52
M12	22,90	2,61
M13	32,53	8,30
M14	7,64	31,14
M15	18,50	5,90
M16	27,72	3,58
M17	26,20	3,47
M18	22,90	2,61

Karena pada Iterasi ke-2 posisi cluster tidak berubah/sama dengan posisi cluster pada iterasi pertama maka proses iterasi dihentikan.

{M1, M2,M3,M4,M5,M6,M14} : Anggota C1

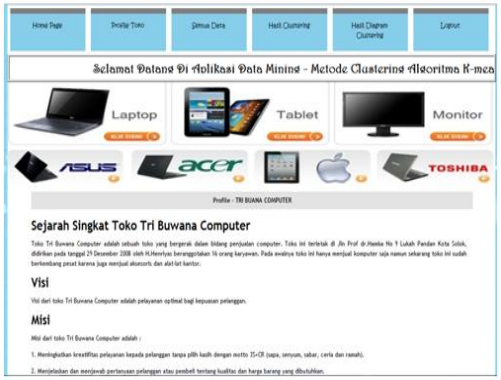
{,M7,M8,M9,M10,M11,M12,M13,M15,M16 ,17,M18} : Anggota C2
Kesimpulan :

M1,M2,M3,M4,M5,M6,M14 merupakan anggota dari cluster 1 dan M7,M8,M9,M10,M11,M12,M15,M16,M17,M 18 merupakan anggota dari cluster 2.

4. PENGUJIAN

1. Tampilan Home

2. Tampilan Profil



3. Tampilan Semua Data

4. Tampilan Hasil Clustering

5. Tampilan Hasil Diagram Clustering

5. PENUTUP

Kesimpulan

Dari penulisan penelitian ini mulai dari tahapan analisa permasalahan yang ada, hingga pengujian aplikasi sistem yang baru maka dapat diambil beberapa kesimpulan yaitu :

1. Aplikasi data mining ini dapat membantu pimpinan untuk mengambil keputusan dalam meningkatkan efektifitas penjualan laptop di toko Tri Buana Computer yang sebelumnya masih cenderung lambat.
2. Aplikasi data mining memberikan kemudahan dalam pengolahan data, pimpinan hanya menginputkan data lalu aplikasi akan melakukan proses dan melakukan hasil analisa.
3. Dari hasil analisa yang telah dilakukan maka aplikasi dapat membantu mempermudah pihak toko dalam memprediksi tingkat penjualan.

Dari hasil analisis yang dilakukan, sistem yang dirancang masih memiliki keterbatasan yaitu, aplikasi data mining ini pada intinya hanya bisa memberikan informasi kepada pimpinan untuk menganalisa dan mengambil keputusan terhadap proses analisa barang

DAFTAR PUSTAKA

Hermawati, Fajar Astuti.2013, *Data Mining*, Bandung : Dunia Koputer.

Kusrini, Emha Taufiq Lutfhi, 2011 Algoritma Data Mining. Bandung : Andi Publisher.

Oktavian, Diar Puji, 2013. Komputerpedia Membuat Website Powerfull Menggunakan PHP. Yogyakarta: MEDIA KOM

Sommerville, Ian, *2013*, *Software Enginering*. Jakarta : Erlangga.

S, Rosa A. dan M Shalahuddin, 2013. *Rekayasa Perangkat Lunak Terstruktur dan Berorientasi Objek.* Bandung : Informatika