

STMIK GI MDP

Program Studi Teknik Informatika
Skripsi Sarjana Komputer
Semester Ganjil Tahun 2011/2012

CLUSTERING **DATA PENJUALAN DAN PERSEDIAAN BARANG** **PADA PT SAYAP MAS UTAMA DENGAN METODE K-MEANS**

Ahmad Afif 2008250031
Reksa Prayudhi Aruan 2008250093

Abstract

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

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 yaitu memberi informasi yang terkandung dari suatu data, sehingga *user* mendapatkan informasi yang belum didapatkan dari 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 pembagian kelompok dari suatu data.

Kata kunci : *Clustering*, *K-Means*, informasi

PENDAHULUAN

Semakin besar suatu perusahaan, semakin lama perusahaan itu berdiri, semakin banyak perusahaan memiliki cabang, tentunya semakin besar data yang dimiliki. Data-data tersebut bisa berupa data konsumen, data pembelian, data penjualan, data penggajian, dan masih banyak lagi. Semua data tersebut biasanya akan tersimpan dalam *Database Center*. Namun banyak perusahaan, bahkan divisi IT yang tidak menyadari betapa berharganya tumpukan data-data lama yang dihasilkan perusahaan dalam bertransaksi dan beraktifitas.

Kebanyakan informasi hanya dipandang sebagai arsip yang sudah menjadi berkas-berkas tidak terpakai dan bisa dihancurkan kapan saja. Hal tersebut tentu saja merupakan pandangan yang salah, sebab dengan penanganan yang tepat dan cerdas data-data tersebut dapat diberdayakan dengan *Data Mining*, sehingga nantinya dapat digunakan untuk meramalkan strategi bisnis dan masa depan perusahaan.

Menambang data atau *data mining* adalah sebuah metode yang digunakan untuk mengekstraksi informasi prediktif tersembunyi pada *database* yang mampu memprediksi tren, perilaku, sehingga memungkinkan perusahaan untuk semakin proaktif dan memperkaya pengetahuan atau informasi dalam membuat keputusan. Pemanfaatan *data mining* dapat dilihat dari 2 (dua) sudut pandang, yaitu sudut pandang komersial dan sudut pandang keilmuan. Dari sudut pandang komersial, pemanfaatan *data mining* dapat digunakan dalam menangani meledaknya volume data. Bagaimana cara menyimpannya, mengekstraknya serta memanfaatkannya. Berbagai teknik komputasi dapat digunakan menghasilkan informasi yang dibutuhkan. Informasi yang dihasilkan menjadi aset untuk meningkatkan daya saing suatu institusi. Sedangkan dari sudut pandang keilmuan, data mining dapat digunakan untuk *capture*, menganalisis serta menyimpan data yang bersifat real-time dan sangat besar.

Perusahaan berskala menengah ke atas seperti PT Sayap Mas Utama saat ini memang sudah berkewajiban untuk memanfaatkan *data mining* jika memang ingin sukses dalam kancah bisnis khususnya didalam bagian penjualan produk-produknya. Jika tidak, manajemen yang dilakukan tidak ada bedanya dengan

perusahaan tradisional yang kemudian beralih menggunakan komputer yang hanya menggantikan cara-cara manual. Perusahaan besar sekalipun akan mengalami kemunduran dalam bisnisnya, jika tidak adanya perencanaan yang baik untuk prospek bisnisnya kedepan, dengan menggunakan *data mining* tentu saja hal ini bisa diatasi.

Salah satu metode yang terdapat dalam *data mining* yang digunakan dalam penelitian ini adalah pengelompokan (*Clustering*) dimana metode tersebut mengidentifikasi objek yang memiliki kesamaan karakteristik tertentu, dan kemudian menggunakan karakteristik tersebut sebagai “vektor karakteristik” atau “*centroid*”. Pengelompokan ini digunakan oleh pihak manajemen perusahaan untuk membuat laporan mengenai karakteristik umum dari grup-grup konsumen yang berbeda. Proses *Clustering* yang akan dilakukan menggunakan Algoritma *K-Means*.

Berdasarkan hal tersebut maka diharapkan bisa mempermudah pihak perusahaan dalam menemukan informasi dari tumpukan-tumpukan data yang berguna bagi perusahaan. Tidak hanya itu penulis juga mengharapkan bisa memberikan rangsangan kepada pengelola data supaya lebih aktif dalam menggali informasi yang dibutuhkan oleh perusahaan, sehingga bisa menjawab kebutuhan bisnis dari perusahaan. Bagi pihak manajemen level atas perusahaan PT Sayap Mas Utama penulis berharap dapat membantu memberikan informasi mengenai hasil *Clustering* data penjualan.

METODOLOGI

A. Analisis dan definisi persyaratan. Pelayanan, batasan, dan tujuan sistem ditentukan melalui konsultasi dengan *user* sistem.

Persyaratan ini kemudian didefinisikan secara rinci dan berfungsi sebagai spesifikasi sistem.

B. Perancangan sistem dan perangkat lunak.

Proses perancangan sistem membagi persyaratan dalam sistem perangkat keras atau perangkat lunak. Kegiatan ini menentukan arsitektur sistem secara keseluruhan.

C. Implementasi dan pengujian unit.

Pada tahap ini, perancangan perangkat lunak direalisasikan sebagai serangkaian program atau unit program. Pengujian unit melibatkan verifikasi bahwa setiap unit memenuhi spesifikasinya.

D. Integrasi dan pengujian sistem.

Unit program atau program individual diintegrasikan dan diuji sebagai sistem yang lengkap untuk menjamin bahwa persyaratan sistem telah dipenuhi.

E. Operasi dan pemeliharaan.

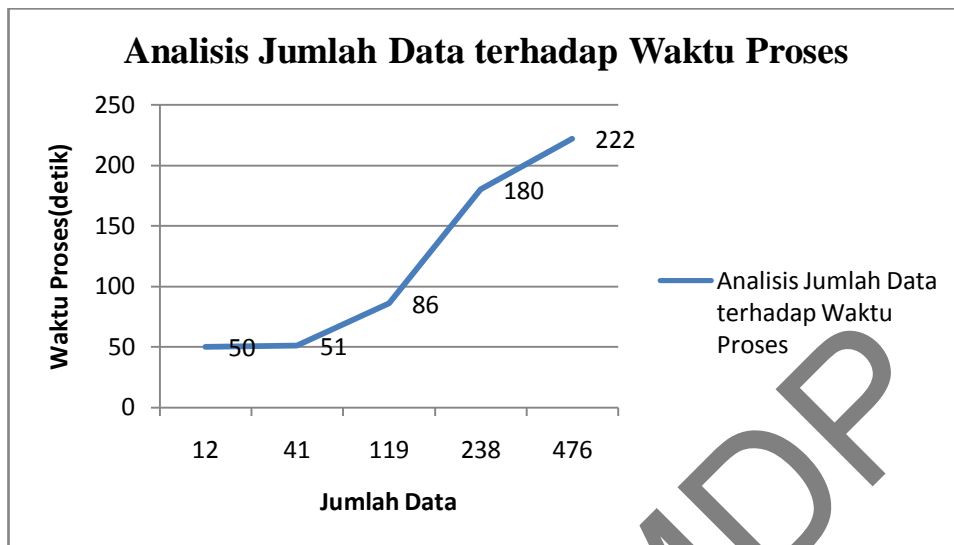
Sistem diinstal dan dipakai. Pemeliharaan mencakup koreksi dari berbagai *error* yang tidak ditemukan pada tahap-tahap terdahulu, perbaikan atas implementasi unit sistem dan pengembangan pelayanan sistem, sementara persyaratan-persyaratan baru ditambahkan.

HASIL DAN PEMBAHASAN

Tabel 1.1 dan tabel 1.2 menunjukkan jumlah data terhadap waktu proses dan jumlah data terhadap banyak iterasi dari 5 kali pengujian dengan 2 (*cluster*).

Tabel 1.1 Analisis Jumlah Data Terhadap Waktu Proses

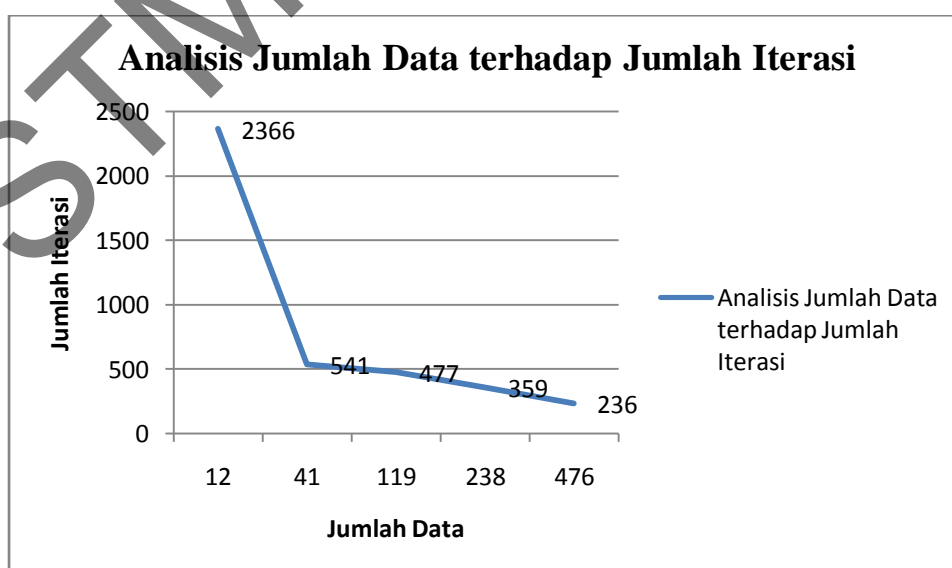
Percobaan	Jumlah Data	Waktu Proses
1	12	± 50 detik
2	41	± 51 detik
3	119	± 1 menit 26 detik
4	238	± 3 Menit
5	476	± 3 menit 42 detik



Gambar 1.1 Grafik Analisis Jumlah Data Terhadap Waktu Proses

Tabel 4.14 Analisis Jumlah Data Terhadap Jumlah Iterasi

Percobaan	Jumlah Data	Banyak Iterasi
1	12	2366
2	41	541
3	119	477
4	238	359
5	476	236



Gambar 4.19 Grafik Analisis Waktu Respon Terhadap Jumlah Iterasi

KESIMPULAN DAN SARAN

Dari hasil pengamatan selama perencanaan, implementasi dan uji coba perangkat lunak, dapat diambil kesimpulan sebagai berikut :

1. Sistem yang dibuat mampu memenuhi tujuan awal pembuatan aplikasi yaitu melakukan penerapan *data mining* pada data penjualan dan persediaan barang/produk di PT Sayap Mas Utama dengan metode *Clustering*.
2. Berdasarkan uji coba, semakin kecil pengelompokkan data (*cluster*) yang dipilih dengan jumlah data yang kecil, semakin cepat juga aplikasi menampilkan pengelompokkan data, spesifikasi komputer/laptop berpengaruh dengan proses penampilan data.
3. Penampilan data yang dikelompokkan (*cluster*) masih bersifat tabel.
4. Pihak manajemen perusahaan dapat mengetahui banyak atau sedikitnya barang yang dibutuhkan dari pengelompokkan data sehingga dapat memberikan keputusan yang tepat mengenai barang-barang tersebut. Barang yang sangat dibutuhkan akan diproduksi dalam jumlah yang besar dan barang yang sedikit dibutuhkan akan diproduksi dalam jumlah yang kecil untuk mengurangi beban kapasitas gudang dan resiko terjadinya kerusakan yang berdampak kerugian pada perusahaan.
5. Berdasarkan hasil kuisioner terhadap dua puluh *user* yang telah dipilih secara acak untuk menilai kinerja dari aplikasi ini dapat disimpulkan bahwa aplikasi *Clustering* data penjualan dan persediaan barang pada PT Sayap Mas Utama dengan metode *K-Means*, tampilan *user interface* nya menarik, penggunaannya cukup mudah, proses *Clustering* dengan metode *K-Means* sudah sesuai dengan kebutuhan, dan informasi hasil proses *Clustering* nya juga sudah cukup memadai sebagai bahan pertimbangan pihak manajemen perusahaan didalam mengambil keputusan mengenai strategi apa yang akan dilakukan. Akan tetapi, aplikasi ini masih memiliki banyak kekurangan, salah satu nya adalah untuk perhitungan *Clustering* data yang jumlahnya banyak, waktu prosesnya relatif cukup lama dibandingkan dengan perhitungan *Clustering* data yang jumlahnya sedikit.

Saran

Saran dalam kemungkinan pengembangan aplikasi yang dilakukan, yaitu :

1. Diharapkan jumlah kelompok (*cluster*) tidak terbatas, sehingga *user* bebas menentukan jumlah kelompok sesuai dengan keinginannya.
2. Diharapkan tampilan data yang dikelompokkan dapat lebih menarik lagi.
3. Diharapkan menggunakan spesifikasi komputer/laptop yang tinggi seperti komputer/laptop yang memiliki *processor* Intel Core i7 dan memiliki RAM diatas 2GB untuk bisa mempercepat proses *Clustering* data penjualannya.

DAFTAR PUSTAKA

Han, Jiawei; Michaline Kamber; dan Jian Pei 2011, *Data Mining Concepts and Techniques*, Morgan Kaufmann Publishers, USA.

Irdiansyah, Enur 2009, *Penerapan Data mining pada Penjualan Produk Minuman Di PT Pepsi Cola Indobeverages Menggunakan Metode Clustering*, Universitas Komputer Indonesia (UNIKOM), Bandung,

Praduwiratna, Rangga 2010, *Bermain Data dengan SQL Server*, Microsoft Student Partners, Bandung.

Schlesinger, Richard 2008, *The Visual basic .Net Programming Language*, Jones and Bartlett Publishers, London.

Sulianta, Feri 2010, *Data mining Meramalkan Bisnis Perusahaan*, Elex Media Komputindo, Jakarta.

Susanto, Sani dan Dedy Suryadi 2010, *Pengantar Data mining*, Andi Offset, Yogyakarta.

Wahana, Komputer 2010, *Membuat Aplikasi Client Server dengan Visual Basic* 2008, Andi Publisher, Yogyakarta.

Lampiran 1

Analisis Cluster K-Means

Cari Data

☐ Tanggal : 31 Januari 2012

☐ Tanggal : 31 Januari 2012 s/d 31 Januari 2012

☒ Pilih Berdasarkan : Daerah

Keterangan Nilai

BCV : 36316.0305925634 Ratio 1 0.0000000029

WCV : 12538017422000 Ratio 2 0.0000000029

No.	Item Code	Item Name	Capacity	Initial Stock Qty	Received Qty
1	20001	SO KLIN FLOOR CLEANER ...	1600 ML	85	95
2	20021	SO KLIN FLOOR CLEANER ...	800 ML	696	556
3	20036	SO KLIN FLOOR CLEANER ...	4000 ML	41	60
4	20063	SO KLIN FLOOR CLEANER ...	400 ML	276	168
5	31061	WINGS PORCELAIN BTL 80...	800 ML	37	90
6	31062	WINGS PORCELAIN BTL 2000	2000 ML	3	9
7	31068	WINGS SUPER SOL POUNC...	800 ML	13	48

Keterangan Cluster

Cluster 1 : 0 - 16392

Cluster 2 : 16392 - 32784

Cluster 3 : 32784 - 49176

Cluster 4 :

Cluster 5 :

Jumlah cluster 3

Jumlah Iterasi : 28

No	Item Code	Item Name	Capacity	Cluster	Keterangan
1	20001	SO KLIN FLOOR CLEANER REFI	1600 ML	Cluster 1	Tidak Baik
2	20021	SO KLIN FLOOR CLEANER POU	800 ML	Cluster 1	Tidak Baik
3	20036	SO KLIN FLOOR CLEANER BTL	4000 ML	Cluster 1	Tidak Baik
4	20063	SO KLIN FLOOR CLEANER POU	400 ML	Cluster 1	Tidak Baik
5	31061	WINGS PORCELAIN BTL 800 M	800 ML	Cluster 1	Tidak Baik
6	31062	WINGS PORCELAIN BTL 2000	2000 ML	Cluster 1	Tidak Baik
7	31068	WINGS SUPER SOL POUNCH 800 M	800 ML	Cluster 1	Tidak Baik
8	31069	WINGS SUPER SOL BTL 900 ML	900 ML	Cluster 1	Tidak Baik
9	31091	WINGS SUPER SOL REFIL 18	1800 ML	Cluster 1	Tidak Baik

Waktu Proses
Awal : 15:30:49 Akhir : 15:31:06 Total (detik) : 17

Petunjuk

1. Tentukan Data yang hendak di cluster dengan cara memilih rentang tanggal yang ada dibagian kiri atas form, lalu klik tombol cari atau anda dapat memilih tombol tampil semua untuk menampilkan semua data.
2. Tentukan berapa banyak jumlah cluster yang akan diproses, banyaknya cluste berjumlah antara 2 sampai 5 cluster.
3. Lalu tekan OK dan Sistem akan memulai proses clustering
4. Jika proses telah selesai, hasil proses nya akan tampak di datagrid disamping.
5. Pilih tombol bersih jika anda ingin mengcluster dengan jumlah cluster yang lain.

Gambar 1 Tampilan Form Analisis Cluster

Lampiran 2

Hasil Clustering Data Penjualan PT Sayap Mas Utama

17/02/2012
12:24:51

ItemCode	ItemName	Capacity	Cluster	Keterangan
20001	SO KLIN FLOOR CLEANER REFI	1600 ML	Cluster 1	Tidak Baik
20021	SO KLIN FLOOR CLEANER POU	800 ML	Cluster 3	Sangat Baik
20036	SO KLIN FLOOR CLEANER BTL	4000 ML	Cluster 1	Tidak Baik
20063	SO KLIN FLOOR CLEANER POU	400 ML	Cluster 2	Baik
20084	SO KLIN FLOOR CLEANER SACH	25 ML x 12	Cluster 1	Tidak Baik
22035	CLING GLASS CLEANER BOTOL	425 ML	Cluster 1	Tidak Baik
22036	CLING GLASS CLEANER REFIL	425 ML	Cluster 2	Baik
31061	WINGS PORCELAIN BTL 800 M	800 ML	Cluster 1	Tidak Baik
31062	WINGS PORCELAIN BTL 2000	2000 ML	Cluster 1	Tidak Baik
31068	WINGS SUPER SOL POUNCH 800	800 ML	Cluster 1	Tidak Baik
31069	WINGS SUPER SOL BTL 900 ML	900 ML	Cluster 1	Tidak Baik
31091	WINGS SUPER SOL REFIL 18	1800 ML	Cluster 1	Tidak Baik

Keterangan Cluster

Cluster 1 : 0 - 887

Cluster 2 : 888 - 1774

Cluster 3 : 1775 - 2662

Gambar 2 Form Cetak Hasil Clustering Data dengan Tiga Cluster