



## Makalah DATA Mining

mathematics (Universitas Muria Kudus)

# **MAKALAH DATA MINING**



Disusun oleh :

Mirza Faqihuddin Ahmad (201953143)

**UNIVERSITAS MURIA KUDUS**  
**FAKULTAS TEKNIK**  
**PROGRAM STUDI SISTEM INFORMASI**  
**2022**

## DAFTAR ISI

BAB I.....	3
PENDAHULUAN.....	3
1.1.    Maksud dan Tujuan Penelitian.....	3
1.2.    Batasan Masalah.....	3
BAB II.....	4
PEMBAHASAN.....	4
2.1.    Pengertian Data Mining.....	4
2.2.    Teknik Data Mining.....	5
2.3.    Implementasi (penerapan) Data Mining.....	6
2.4.    Contoh Kasus Penerapan.....	7
2.5.    Implementasi dan Hasil Analisa.....	9
BAB III.....	12
PENUTUP.....	12
3.1    Kesimpulan.....	12
3.1    Saran.....	13

# **BAB I**

## **PENDAHULUAN**

### **1.1. Maksud dan Tujuan Penelitian**

- Menerapkan proses data mining untuk pengolahan basis data customer dengan metode clustering menggunakan algoritma hirarkis devisie k-means untuk mengelompokkan customer.
- Penggunaan metode clustering dan algoritma hirarki divisive k-means untuk mengetahui kemiripan karakteristik antar data dalam basis data customer berdasarkan transaksi yang dilakukan, guna membentuk kelompok-kelompok customer.
- Membangun sebuah aplikasi data mining yang dapat mentransformasikan basis data customer berdasarkan transaksi yang dilakukan menjadi informasi yang berguna.
- Memanfaatkan aplikasi data mining untuk melakukan competitive intelligence perusahaan guna pengelompokan customer.

### **1.2. Batasan Masalah**

Batas an perumusan masalah dalam penulisan studi kasus pada tugas ini, meliputi:

- Competitive intelligence dalam studi kasus ini hanya sebatas untuk melakukan pengelompokan customer berdasarkan data transaksi yang dilakukan saja tanpa melakukan proses competitive intelligence lainnya.
- Basis data yang akan digunakan dalam studi kasus ini adalah basis data customer dan transaksi yang dilakukan saja tanpa melibatkan basis data lainnya, yang kemudian akandiolah berdasarkan proses-proses yang ada dalam data mining.
- Kemiripan antar data dalam studi kasus ini diterjemahkan sebagai jarak kedekatan antar data dengan titik pusat, sehingga menghasi lkan klaster-klaster customer yang sesuai dengan tujuan dari studi kasus ini.
- Penggunaan metode clustering untuk mengelompokan customer dengan menggunakan algoritma hirarki divisive k-means.

## **BAB II**

### **PEMBAHASAN**

#### **IMPLEMENTASI DATA MINING DENGAN METODE CLUSTERING UNTUK MELAKUKAN COMPETITIVE INTELLIGENCE PERUSAHAAN**

##### **2.1. Pengertian Data Mining**

Data Mining adalah serangkaian proses untuk menggali nilai tambah dari suatu kumpulan data berupa pengetahuan yang selama ini tidak diketahui secara manual. Patut diingat bahwa kata mining sendiri berarti usaha untuk mendapatkan sedikit barang berharga dari sejumlah besar material dasar. Karena itu DM sebenarnya memiliki akar yang panjang dari bidang ilmu seperti kecerdasan buatan (artificial intelligent), machine learning, statistik dan database. Data mining adalah proses menerapkan metode ini untuk data dengan maksud untuk mengungkap pola-pola tersembunyi. Dengan arti lain Data mining adalah proses untuk penggalan pola-pola dari data. Data mining menjadi alat yang semakin penting untuk mengubah data tersebut menjadi informasi. Hal ini sering digunakan dalam berbagai praktek profil, seperti pemasaran, pengawasan, penipuan deteksi dan penemuan ilmiah. Telah digunakan selama bertahun-tahun oleh bisnis, ilmuwan dan pemerintah untuk menyaring volume data seperti catatan perjalanan penumpang penerbangan, data sensus dan supermarket scanner data untuk menghasilkan laporan riset pasar.

Alasan utama untuk menggunakan data mining adalah untuk membantu dalam analisis koleksi pengamatan perilaku. Data tersebut rentan terhadap collinearity karena diketahui keterkaitan. Fakta yang tak terelakkan data mining adalah bahwa subset/set data yang dianalisis mungkin tidak mewakili seluruh domain, dan karenanya tidak boleh berisi contoh-contoh hubungan kritis tertentu dan perilaku yang ada di bagian lain dari domain. Untuk mengatasi masalah semacam ini, analisis dapat ditambah menggunakan berbasis percobaan dan pendekatan lain, seperti Choice Modelling untuk data yang dihasilkan manusia.

Dalam situasi ini, yang melekat dapat berupa korelasi dikontrol untuk, atau dihapus sama sekali, selama konstruksi desain eksperimental. Beberapa teknik yang sering disebut-sebut dalam literatur Data Mining dalam penerapannya antara lain: clustering, classification, association rule mining, neural network, genetic algorithm dan lain-lain. Yang membedakan persepsi terhadap Data Mining adalah perkembangan teknik-teknik Data Mining untuk aplikasi 1 pada database skala besar. Sebelum populernya Data Mining, teknik-teknik tersebut hanya dapat dipakai untuk data skala kecil saja.

##### **Proses Data Mining**

Tahap-Tahap Data Mining. Karena Data Mining adalah suatu rangkaian proses, Data Mining dapat dibagi menjadi beberapa tahap :

1. Pembersihan data (untuk membuang data yang tidak konsisten dan noise)
2. Integrasi data (penggabungan data dari beberapa sumber)
3. Transformasi data (data diubah menjadi bentuk yang sesuai untuk di-mining)
4. Aplikasi teknik Data Mining
5. Evaluasi pola yang ditemukan (untuk menemukan yang menarik/bernilai)
6. Presentasi pengetahuan (dengan teknik visualisasi)

## 2.2. Teknik Data Mining

Berikut beberapa jenis teknik Data Mining yang paling populer dikenal dan digunakan:

### 1. Association Rule Mining

Association rule mining adalah teknik mining untuk menemukan aturan assosiatif antara suatu kombinasi item. Penting tidaknya suatu aturan assosiatif dapat diketahui dengan dua parameter, support yaitu persentase kombinasi item tsb. dalam database dan confidence yaitu kuatnya hubungan antar item dalam aturan assosiatif. Algoritma yang paling populer dikenal sebagai Apriori dengan paradigma generate and test, yaitu pembuatan kandidat kombinasi item yang mungkin berdasar aturan tertentu lalu diuji apakah kombinasi item tsb memenuhi syarat support minimum. Kombinasi item yang memenuhi syarat tsb. disebut frequent itemset, yang nantinya dipakai untuk membuat aturan-aturan yang memenuhi syarat confidence minimum. Algoritma baru yang lebih efisien bernama FP-Tree.

### 2. Classification Classification

Classification Classification adalah proses untuk menemukan model atau fungsi yang menjelaskan atau membedakan konsep atau kelas data, dengan tujuan untuk dapat memperkirakan kelas dari suatu objek yang labelnya tidak diketahui. Model itu sendiri bisa berupa aturan “jika-maka”, berupa decision tree, formula matematis atau neural network. Decision tree adalah salah satu metode classification yang paling populer karena mudah untuk diinterpretasi oleh manusia. Disini setiap percabangan menyatakan kondisi yang harus dipenuhi dan tiap ujung pohon menyatakan kelas data.

Algoritma decision tree yang paling terkenal adalah C4.5, tetapi akhir-akhir ini telah dikembangkan algoritma yang mampu menangani data skala besar yang tidak dapat ditampung di main memory seperti RainForest. Metode-metode classification yang lain adalah Bayesian, neural network, genetic algorithm, fuzzy, case-based reasoning, dan k-nearest neighbor. Proses classification biasanya dibagi menjadi dua fase : learning dan test. Pada fase learning, sebagian data yang telah diketahui kelas datanya diumpukan untuk membentuk model perkiraan. Kemudian pada fase test model yang sudah terbentuk diuji dengan sebagian data lainnya untuk mengetahui akurasi dari model tsb. Bila akurasinya mencukupi model ini dapat dipakai untuk prediksi kelas data yang belum diketahui.

### 3. Clustering

Berbeda dengan association rule mining dan classification dimana kelas data telah ditentukan sebelumnya, clustering melakukan penge-lompokan data tanpa berdasarkan kelas data tertentu. Bahkan clustering dapat dipakai untuk memberikan label pada kelas data yang belum diketahui itu. Karena itu clustering sering digolongkan sebagai metode unsupervised learning. Prinsip dari clustering adalah memaksimalkan kesamaan antar anggota satu kelas dan meminimumkan kesamaan antar kelas/cluster. Clustering dapat dilakukan pada data yang memiliki beberapa atribut yang dipetakan sebagai ruang multidimensi. Banyak algoritma clustering

memerlukan fungsi jarak untuk mengukur kemiripan antar data, diperlukan juga metode untuk normalisasi bermacam atribut yang dimiliki data.

Beberapa kategori algoritma clustering yang banyak dikenal adalah metode partisi dimana pemakai harus menentukan jumlah ke partisi yang diinginkan lalu setiap data dites untuk dimasukkan pada salah satu partisi, metode lain yang telah lama dikenal adalah metode hierarki yang terbagi dua lagi : bottom-up yang menggabungkan cluster kecil menjadi cluster lebih besar dan top-down yang memecah cluster besar menjadi cluster yang lebih kecil.

Kelemahan 3 metode ini adalah bila salah satu penggabungan/pemecahan dilakukan pada tempat yang salah, tidak dapat didapatkan cluster yang optimal. Pendekatan yang banyak diambil adalah menggabungkan metode hierarki dengan metode clustering lainnya seperti yang dilakukan oleh Chameleon. Akhir-akhir ini dikembangkan juga metode berdasar kepadatan data, yaitu jumlah data yang ada di sekitar suatu data yang sudah teridentifikasi dalam suatu cluster. Bila jumlah data dalam jangkauan tertentu lebih besar dari nilai ambang batas, data-data tersebut dimasukkan dalam cluster. Kelebihan metode ini adalah bentuk cluster yang lebih fleksibel. Algoritma yang terkenal adalah DBSCAN.

#### 4. Algoritma Hirarki Divisive

Langkah awal yang dilakukan dalam algoritma hirarki divisive adalah membentuk satu cluster besar yang dapat ditempati oleh semua obyek data. Pada langkah berikutnya, satu cluster besar tersebut dipisah – pisahkan menjadi beberapa cluster yang lebih kecil dengan karakteristik data yang mempunyai lebih besar kesamaan satu dengan yang lainnya, sehingga data yang tidak memiliki kemiripan yang cukup besar berada pada cluster yang terpisah.

### 2.3. Implementasi (penerapan) Data Mining

Dalam bidang apa saja data mining dapat diterapkan? Berikut beberapa contoh bidang penerapan data mining:

Analisa pasar dan manajemen

Solusi yang dapat diselesaikan dengan data mining, diantaranya: Menembak target pasar, Melihat pola beli pemakai dari waktu ke waktu, Cross-Market analysis, Profil Customer, Identifikasi kebutuhan Customer, Menilai loyalitas Customer, Informasi Summary.

- Analisa Perusahaan dan Manajemen resiko.

Solusi yang dapat diselesaikan dengan data mining, diantaranya: Perencanaan keuangan dan Evaluasi aset, Perencanaan sumber daya (Resource Planning), Persaingan (Competition).

- Telekomunikasi.

Sebuah perusahaan telekomunikasi menerapkan data mining untuk melihat dari jutaan transaksi yang masuk, transaksi mana sajakah yang masih harus ditangani secara manual.

- Keuangan.

Financial Crimes Enforcement Network di Amerika Serikat baru-baru ini menggunakan data mining untuk me-nambang trilyunan dari berbagai subyek seperti property, rekening bank dan transaksi keuangan lainnya untuk mendeteksi transaksi-transaksi keuangan yang mencurigakan (seperti money laundry) .

- Asuransi.

Financial Crimes Enforcement Network di Amerika Serikat baru-baru ini menggunakan data mining untuk me-nambang trilyunan dari berbagai subyek seperti property, rekening bank dan transaksi keuangan lainnya untuk mendeteksi transaksi-transaksi keuangan yang mencurigakan (seperti money laundry) .

- Olahraga.

IBM Advanced Scout menggunakan data mining untuk menganalisis statistik permainan NBA (jumlah shots blocked, assists dan fouls) dalam rangka mencapai keunggulan bersaing (competitive advantage) untuk tim New York Knicks dan Miami Heat.

## 2.4. Contoh Kasus Penerapan

“Implementasi data mining dengan teknik Clustering untuk melakukan Competitive Intelligence perusahaan” Pembangunan perangkat lunak data mining dengan metode clustering menggunakan algoritma hirarki divisive untuk pengelompokan customer dalam studi kasus ini, fungsi – fungsi yang dipakai adalah fungsi untuk menentukan titik-titik pusat yang berguna sebagai pusat-pusat kelompok customer.

### 1. Analisa Permasalahan :

Dalam kegiatan bisnis untuk mempertahankan area pemasarannya, toko Benteng jewellry mengalami beberapa permasalahan-permasalahan yang menyangkut kebutuhan data dan informasi tentang customer , sehingga untuk melakukan kegiatan-kegiatan promosi demi mempertahankan pangsa pasar agar tetap bisa bertahan ditengah krisis ekonomi mengalami beberapa masalah. Permasalahan-permasalahan itu antara lain, yaitu :

- Sulitnya melakukan analisa pemasaran yang efektif karena tidak adanya sistem yang dapat menyajikan data historis sehingga dapat memberikan output tentang berapa banyak jumlah customer yang dimiliki dan kelompok-kelompok customer yang aktif maupun tidak menurut frekuensi transaksinya, karena data yang ada masih berbentuk data manual dan belum dimanfaatkan secara maksimal.
- Tidak diketahui dengan pasti jumlah customer yang aktif dan yang kurang aktif dalam melakukan transaksi, sehingga sangat sulit untuk melakukan tindakan- tindakan promosi maupun pemberian bonus atau diskon kepada setiap customer yang dimiliki dengan tepat.
- Terlalu banyaknya competitor usaha, sehingga diperlukan sebuah sistem yang bisa mendeteksi berapa jumlah customer yang aktif dan yang kurang aktif dalam bertransaksi sebagai sistem pendukung keputusan, sehingga bisa digunakan untuk merancang sebuah strategi bisnis yang efektif guna mempertahankan pangsa pasar dalam persaingan dengan competitor ditengah krisis ekonomi global seperti sekarang ini.

### 2. Pemecahan Masalah :



Berdasarkan latar belakang masalah diatas, maka dibutuhkan sebuah system yang mampu mengelola data customer yang dapat memberikan output berupa jumlah customer secara keseluruhan dan kelompok-kelompok customer yang menyatakan keaktifan melakukan transaksi sehingga bisa digunakan untuk melakukan customer relationship guna kelancaran kegiatan promosi untuk mempertahankan pangsa pasar sehingga toko ini bisa bertahan ditengah krisis ekonomi global. Atas dasar analisis diatas, maka penulis tertarik untuk meneliti bidang ini dengan mengambil judul “ Implementasi Data Mining dengan Metode Clustering Untuk Melakukan Competitive Intelligence Perusahaan” guna pengelompokan customer .

#### a. Analisa Kebutuhan Sistem

Analisa kebutuhan sistem berfungsi untuk mendefinisikan kebutuhan-kebutuhan sistem yang akan dibangun. Analisa ini bertujuan untuk menghasilkan data yang bisa diintegrasikan dengan analisa data mining yang dikehendaki.

#### b. Analisa Kebutuhan Data

Analisa data akan mengidentifikasikan kebutuhan data yang sesuai dengan ketentuan yang diperlukan sistem dari data yang tidak lengkap dan inkonsisten yang biasanya terjadi pada basis data yang ada. Analisa ini meliputi:

- Proses data cleaning (pembersihan data).
- Analisa target data.
- Proses data integration (integrasi data).
- Proses data selection (pemilihan data).
- Proses data transformasi (pembentukan data prosesing).
- Analisa kebutuhan data input, proses dan output.

#### c. Analisa Kebutuhan Perangkat Keras Dan Perangkat Lunak

Analisa ini mendeskripsikan perangkat yang dibutuhkan dalam pembangunan sistem yang terdiri dari komponen perangkat keras dan perangkat lunak. Komponen perangkat keras yang dibutuhkan oleh sistem adalah sebuah pc atau workstation dengan spesifikasi minimal, sebagai berikut:

- Hardware

Processor intel Pentium IV atau lebih, RAM 512 atau lebih, HDD 80GB, VGA 12 MB shared, CD-RW/ DVD-RW

- Software

- 111111111 1111 Operating system : windows 98/2000/XP
- XAMPP-win32-1.6.7
- web browser : Ms. Internet Explorer , Mozilla Firefox 3.0

### 3. Perancangan system

Dalam perancangan sistem ini, metode yang digunakan adalah metode perancangan sistem berorientasi objek dengan ( Object Oriented Analysis) dengan menggunakan Unified Modelling Language (UML).

## 2.5. Implementasi dan Hasil Analisa

Dalam bab implementasi dan analisa hasil, akan dijelaskan tentang pembangunan perangkat lunak yang telah dirancang sebelumnya pada bab sebelumnya yaitu bab tentang analisa dan perancangan. Implementasi dari perancangan pembangunan perangkat lunak dalam studi kasus ini, meliputi:

### 1. Implementasi basis data

Dalam studi kasus ini, sistem basis data yang digunakan adalah ApacheFriends XAMPP versi 1.6.7. Karena basis data awal dalam studi kasus ini yang dimiliki berupa basis data dalam bentuk manual, maka pembentukan basis data dalam studi kasus ini dibuat melalui pembangunan basis data baru dengan melakukan pembentukan tiap tabel yang diperlukan dan pengisian data dengan cara input data satu persatu kedalam sistem basis data yang ada diaplikasi ApacheFriends XAMPP bukan melalui proses export dan import data.

Adapun proses pembangunan basis data yang sesuai dengan perangkat lunak yang akan dibangun adalah sebagai berikut:

- Pembentukan tabel – tabel yang dibutuhkan dalam basis data dan menentukan struktur tabelnya.
- Melakukan pembentukan basis data baru, yaitu basis data clustering.
- Melakukan penginputan data kedalam setiap tabel yang ada dalam basis data clustering ini kecuali tabel frekuensi yang merupakan tabel proses dalam studi ini.
- Melakukan pengisian tabel frekuensi yang merupakan data proses dalam studi kasus ini, dengan melakukan pembacaan id\_customer yang ada dalam tabel transaksi berdasarkan id\_customer yang ada ditabel customer yang dilakukan oleh sistem secara langsung.

### 2. Implementasi Fungsi

Pembangunan perangkat lunak data mining dengan metode clustering menggunakan algoritma hirarki divisive untuk pengelompokan customer dalam studi kasus ini, fungsi – fungsi yang dipakai adalah fungsi untuk menentukan titik-titik pusat yang berguna sebagai pusat-pusat kelompok customer. Fungsi - fungsi tersebut adalah sebagai berikut (santosa, 2007):

Langkah 1 :

Fungsi untuk menentukan titik pusat awal dari semua data customer yang ada berdasarkan transaksi yang dilakukan menggunakan perhitungan nilai rata-rata (mean) dari semua data yang ada dalam tabel frekuensi transaksi. Pada langkah ini digunakan perhitungan nilai rata-rata (mean) karena untuk mengantisipasi adanya nilai outlier (nilai yang letaknya sangat jauh dari data yang ada) dari data yang ada dalam tabel frekuensi.

Langkah 2 :

Setelah klaster-klaster pada langkah 1 terbentuk, maka pada langkah 2 ini dilakukan pengecekan ulang perhitungan titik pusat setiap klaster dengan menggunakan perhitungan nilai median (perhitungan nilai tengah). Pemakaian perhitungan nilai median karena semua data yang ada sudah diketahui pada langkah 1, sehingga tidak ada kekhawatiran munculnya data outline.

### Langkah 3 :

Fungsi yang digunakan pada langkah ini, sama seperti fungsi yang digunakan pada langkah II, yaitu penggunaan perhitungan nilai tengah (median). Fungsi dalam langkah ini digunakan untuk mengecek apakah titik pusat klaster yang telah terbentuk pada langkah sebelumnya sudah tidak berubah lagi atau tidak, dengan cara membandingkan hasil perhitungan titik pusat langkah ini dengan langkah sebelumnya. Apabila titik pusat tersebut sudah tidak berubah maka pembentukan klaster customer sudah selesai. Tetapi apa bila titik pusat masih berubah maka dilakukan perhitungan ulang seperti pada langkah II, perhitungan ini akan terus berulang sampai titik pusat cluster tidak berubah lagi.

### 3. Implementasi Sistem

Dalam studi kasus ini, sistem yang dibangun merupakan sebuah perangkat lunak data mining dengan metode clustering menggunakan algoritma hirarki divisive. Perangkat lunak ini berisi form-form tampilan basis data yang telah dinormalisasi, form tampilan untuk data frekuensi transaksi dan form tampilan dari hasil pengelompokan data customer menjadi beberapa klaster. Perangkat lunak ini berfungsi untuk mencari pola-pola yang menarik dari basis data yang berupa nilai frekuensi transaksi untuk mengelompokkan customer.

Software yang digunakan untuk membangun perangkat lunak ini adalah kode program PHP dengan server basis data XAMPP. Untuk menjalankannya, cukup dengan menggunakan aplikasi web browser seperti windows internet explorer, firefox, flock web browser, atau aplikasi web browser sejenisnya tetapi harus dengan aplikasi server basis data yaitu XAMPP sudah terinstall pada PC tempat membuka aplikasi ini, hal ini disebabkan karena perangkat lunak ini membutuhkan asupan basis data dalam bentuk tabel untuk bisa melakukan proses clustering. Hasil akhir dari perangkat lunak yang dibangun dalam studi kasus ini berupa tabel pengelompokan customer dan chart-chart presentase dari tabel pengelompokan tersebut, sehingga dari hasil ini nantinya bisa digunakan oleh user dalam hal ini adalah manager pemasaran dan customer service sebagai pendukung keputusan seperti peningkatan promosi kepada customer-customer yang kurang aktif, ataupun strategi-strategi bisnis lainnya.

Adapun form-form atau halaman-halaman utama yang menjadi isi dari perangkat lunak atau aplikasi ini, adalah sebagai berikut:

- Halaman login.
- Halaman menu utama dan view data normal.
- Halaman view data frekuensi.
- Halaman klaster

#### 4. Pengujian Sistem

Pengujian sistem digunakan untuk mengecek performan sistem ketika seorang user dalam hal ini manager menjalankan sistem Pengujian ini, meliputi:

- Pengujian hak akses atau login.
- Pengujian pembentukan klaster berdasarkan data.

## **BAB III**

### **PENUTUP**

#### **.1 Kesimpulan**

Berdasarkan studi kasus yang telah dilakukan, mulai dari tahap studi pustaka, studi observasi, perancangan dan implementasi sistem, maka dari hasil yang didapatkan bisa ditarik kesimpulan sebagai berikut:

- 1) Metode clustering dengan algoritma hirarki divisive bisa digunakan untuk melakukan pengelompokan customer guna competitive intellegent bisnis perusahaan.
- 2) Informasi dari frekuensi transaksi seorang customer bisa digunakan untuk membangun sebuah system yang dapat mentransformasikan data customer menjadi informasi yang berguna untuk melakukan proses competitive intellegent bisnis perusahaan.
- 3) Program hanya dirancang untuk satu user saja yaitu manager pemasaran dan customer service, sehingga user lain yang tidak sah tidak akan bisa melakukan akses keprogram ini karena username dan password sebagai hak akses sah program hanya dirancang untuk satu user saja dan tidak disediakan fasilitas untuk penambahan hak akses.
- 4) Program akan melakukan proses peng-update-an secara otomatis apabila ada perubahan pada basis data terutama pada tabel transaksi dan tabel customer terhadap nilai-nilai data yang ada di tabel frekuensi transaksi dan juga tabel – tabel hasil proses cluster serta grafik – grafik persentasenya.
- 5) Untuk melakukan pembentukan sebuah cluster dibutuhkan sebuah titik pusat yang bisa dicari dari seluruh data yang ada dalam tabel frekuensi transaksi dengan melakukan metode perhitungan nilai rata rata (mean) ataupun perhitungan nilai tengah ( median).
- 6) Aplikasi tetap bisa berjalan baik ketika dijalankan pada tiga web browser yang berbeda, yaitu internet explorer, mozilla firefox dan flock web browser.
- 7) Hasil dari aplikasi ini bisa dijadikan sebagai pendukung keputusan oleh manager terhadap customer – customer yang dimilikiny a. Misalnya pe ndukung keputusan untuk men ingkatkan promosi kebeberapa customer yang berada di klaster kurang aktif dan sedang ataupun keputusan untuk memberikan fasilitas yang lebih exclusive maupun pemberian bonus atau diskon kepada customer yang berada di klaster yang aktif.

## .1 Saran

Berikut adalah saran yang mungkin perlu dilakukan dalam pengembangan selanjutnya dimasa yang akan datang terhadap aplikasi data mining dengan metode clustering menggunakan algoritma hirarki divisive k-means ini nantinya.

- 1) Dimasa yang akan datang dalam pengembangan selanjutnya, guna lebih memaksimalkan pendukung keputusan yang akan diambil, misalkan untuk kepentingan memudahkan kegiatan promosi bisa ditambahkan sebuah fasilitas berupa fasilitas pengiriman email kepada customer .
- 2) Dalam studi kasus ini, item yang digunakan sebagai data proses dalam pembentuk sebuah cluster hanya didasarkan pada satu item saja yaitu pembacaan frekuensi id customer yang ada ditabel transaksi berdasarkan id customer yang ada ditabel customer . Pada pengembangan selanjutnya disarankan untuk pengambilan data proses tidak hanya berdasarkan satu item saja, mungkin juga bisa dilakukan dengan pembacaan lebih dari satu item. Misalnya id barang atau total harga yang dibayarkan untuk transaksinya, sehingga bisa diketahui barang apa saja yang biasa dibeli customer pada sebuah klaster maupun besarnya jumlah total harga yang dibayarkan oleh customer terhadap transaksi yang dilakukannya. Dengan demikian klaster yang dibentuk tidak hanya tiga klaster tapi mungkin lebih dari pada itu dan informasi yang didapatkanpun menjadi lebih banyak.
- 3) Dalam studi kasus ini, algoritma yang digunakan adalah algoritma hirarki divisive yang berfungsi untuk membagi sebuah cluster besar menjadi beberapa cluster kecil, dalam pengembangan selanjutnya dimungkinkan menggunakan metode yang lain, seperti :
  - Metode market basket analisis sehingga bisa diketahui pola-pola lain seperti barang apa saja yang menjadi favorit customer dalam sebuah cluster.
  - Penggunaan algoritma apriori untuk menganalisa kecenderungan seorang customer dalam melakukan transaksi, misalnya untuk meneliti tanggal berapakah biasanya seorang customer melakukan transaksi pembelian dan barang apa saja yang biasa mereka beli sehingga manager bisa menyusun strategi lain dalam pemasarannya.