



Nama : Eriska Febrianto
NIM : 312010617
Kelas : TI.19.C.1
Mata Kuliah : Data Mining

TUGAS 1

Soal :

1. Jelaskan definisi mengenai data mining menurut pendapat para ahli !
2. Sebutkan 7 fungsi dasar dari data mining !
3. Jelaskan kegunaan dari cluster analysis dalam data mining !
4. Sebutkan dan jelaskan tahapan dari proses data mining !
5. Sebutkan empat penerapan data mining dalam berbagai sector !

Jawab :

1. "Data Mining merupakan bagian proses dari KDD (Knowledge Discovery in Databases) yang memiliki beberapa tahapan seperti pemilihan data, pra pengolahan, transformasi, data mining, dan evaluasi hasil. KDD secara umum disebut juga pangkalan data". (Maimon dan Last, 2000). Sedangkan menurut (Turban et al, 2005) menegaskan, "*Data Mining* adalah proses semi otomatis menggunakan Teknik statistic, matematika, kecerdasan buatan, dan *machine learning* untuk mengekstraksi dan mengidentifikasi informasi pengetahuan potensial dan berguna serta tersimpan pada database besar". Sama halnya dengan Maimon dan Last, Data mining menurut (Santoso, 2007) mengemukakan bahwa, "*Data Mining* disebut juga *Knowledge Discovery in Database* (KDD). KDD merupakan kegiatan yang meliputi pengumpulan, pemakaian data, hubungan ataupun pola dalam set data berukuran besar, dan historis menemukan ketaraturan". Berbeda dengan pendapat (Pramudiono, 2007), "*Data Mining* sering disebut juga serangkaian proses untuk menggali nilai tambah berupa pengetahuan yang selama ini tidak diketahui secara manual pada suatu kumpulan dari data".
2. Terdapat 7 fungsi dasar dari Data Mining, diantaranya :
 - 1) **Prediction atau prediksi**, adalah proses untuk menemukan pola yang hilang dari sebuah data dengan menggunakan analisa regresi untuk menemukan data yang masih tidak tersedia.



- 2) **Sequencing atau sekuensi**, adalah bentuk jamak dari asosiasi yaitu berfungsi sebagai proses untuk menentukan suatu hubungan yang berbeda dari beberapa atau sekumpulan data yang diperoleh.
 - 3) **Classification atau klasifikasi**, mempunyai fungsi menyimpulkan beberapa definisi karakteristik dari suatu grup atau kelompok. Atau bisa dikatakan suatu tindakan untuk memberikan kelompok pada setiap keadaan.
 - 4) **Association atau asosiasi**, adalah identifikasi hubungan dari kejadian-kejadian yang telah terjadi disuatu waktu dengan menyimpulkan definisi karakteristik sebuah data.
 - 5) **Clustering atau pengelompokan**, biasa disebut juga segmentasi (segmentation) berguna dalam proses identifikasi kelompok alami dari sebuah kasus dengan cara mengelompokan data yang memiliki kemiripan.
 - 6) **Forecasting atau perkiraan**, adalah salah satu fungsi dari Data Mining yang bertujuan memperkirakan hasil dari sebuah data di masa yang akan datang berdasarkan pola-pola dengan kumpulan data dalam jumlah besar.
 - 7) **Description atau deskripsi**, adalah memahami lebih jauh karakteristik utama dari suatu data dan membantu menemukan pola tertentu yang masih tersembunyi dalam sebuah data.
3. Kegunaan Cluster Analysis adalah untuk menggambarkan atau visualisasi kelompok objek yang mirip satu sama lain dalam satu kelompok, tetapi berbeda dengan objek dikelompok lain.
4. Proses atau tahapan dalam Data Mining :
- 1) **Data Cleansing**, data yang tidak lengkap atau yang masih kotor dan mengandung error akan dibuang dari koleksi data.
 - 2) **Data Integration**, proses menggabungkan data dari sumber yang berbeda menjadi satu kesatuan.
 - 3) **Data Selection**, proses pemilihan data yang relevan untuk analisis dan didapat dari koleksi data yang ada.
 - 4) **Data Transformation**, proses transformasi data yang terpilih ke dalam bentuk *mining procedure* melalui cara agresi data dan dikonsolidasikan ke bentuk lain sehingga data memiliki distribusi yang diharapkan.



- 5) **Data Mining**, proses yang dilakukan dengan berbagai Teknik yang diaplikasikan untuk mengekstrak pola-pola yang potensial untuk mendapatkan sejumlah data yang berguna.
 - 6) **Pattern Evolution**, Suatu proses dimana informasi menarik yang ditemukan sebelumnya dengan cara identifikasi berdasarkan pengukuran yang telah diberikan.
 - 7) **Knowledge Presentation**, merupakan proses tahap akhir dalam proses data mining, dimana dalam proses ini digunakan teknik visualisasi yang bertujuan untuk membantu pengguna dalam menginterpretasikan hasil dari pencarian data.
5. Empat (4) penerapan Data Mining dalam berbagai sektor, diantaranya :
- 1) **Sektor Telekomunikasi**, Jutaan transaksi yang masuk pada perusahaan telekomunikasi membuat pengumpulan data menerapkan data mining, sehingga dapat diketahui transaksi mana saja yang ditangani secara manual.
 - 2) **Sektor Keuangan**, Data Mining diterapkan pada sektor keuangan yang dilakukan oleh berbagai bank untuk mengumpulkan informasi rekening bank nasabah dan transaksi keuangan lainnya.
 - 3) **Sektor Pendidikan**, Penerapan Data Mining juga dapat dilakukan pada sektor Pendidikan terutama pada proses internal assessment dari masing-masing siswa sekolah dan melakukan prediksi dari hasil ujian nasional siswa.
 - 4) **Sektor Pertanian**, pada sektor ini, Data Mining dapat diterapkan untuk proses pengelompokkan provinsi di Indonesia berdasarkan hasil panen dalam mewujudkan ketahanan pangan nasional.