

Nama : Vinna Jardyagustin
NIM : 19.01.013.064
Kelas : Sains Data 3

Tugas Individu

1. Yang di maksud dengan data mining adalah sekumpulan data berisi fakta dalam jumlah dan ukuran yang besar, data - data tersebut di ubah menjadi satu ilmu pengetahuan yang bermanfaat bagi manusia. Dalam hal ini data - data akan terus bertambah seiring waktu dan kebutuhan penggunaan serta produksi data manusia.

2. Konsep alur proses data mining :

- a. Pembersihan data (pembuangan data yang tidak konsisten atau noise)
- b. Integrasi data (Penggabungan data dari beberapa sumber)
- c. Transformasi data (data di ubah menjadi bentuk sesuai di mining)
- d. Aplikasi teknik data mining proses ekstraksi pola dari data yang ada
- e. Evaluasi pola yang ditemukan (proses interpretasi pola menjadi pengetahuan yang dapat digunakan untuk mendukung pengambilan keputusan)
- f. Presentasi pengetahuan (dengan menggunakan teknik Visualisasi)

3. Peran data mining ada 5, yaitu:

- a. Description
- b. Classification
- c. Prediction
- d. Estimation
- e. Clustering
- f. Association

4. Perbedaan estimation dan forecasting

Estimation (estimasi) merupakan penerkaan atau perkiraan sebuah nilai yang belum di ketahui. Misalnya memperkirakan waktu tempu perjalanan seorang pegawai saat pergi ke kantor, hal itu dapat di prediksi atau diestimasi ketika informasi mengenai orang tersebut di ketahui.

Sedangkan Forecasting (prediksi) adalah untuk menerka nilai masa mendatang. Misalnya, prediksi stok barang sebuah mini market dalam satu bulan ke depan.

5. Perbedaan Forecasting dan klasifikasi
Forecasting (prediksi) adalah menerka nilai masa depan atau masa mendatang. Sedangkan klasifikasi merupakan metode penganalisa data yang sering di masukkan sebagai salah satu metode data mining.

6. Perbedaan klasifikasi dan klastering
Klasifikasi adalah pemrosesan untuk menemukan sebuah model atau fungsi yang menjelaskan dan mencirikan konsep atau kelas data untuk kepentingan tertentu. Sedangkan klastering merupakan pengelompokan data berdasarkan kemiripan pada objek data dan sebaliknya meminimalkan kemiripan terhadap kluster yang lain

7. Perbedaan klostering dan Asosiasi
klostering merupakan pengelompokan data berdasarkan kemiripan objek data. Sedangkan Asosiasi merupakan metode data mining yang di gunakan secara luas dan mengarah ke penemuan yang menarik berhubungan antara Variabel - Variabel

8. Perbedaan estimasi dan klasifikasi
Estimasi adalah untuk menerka sebuah nilai yang belum di ketahui. Sedangkan klasifikasi adalah metode penganalisa data yang sering di masukkan sebagai salah satu metode data mining.

9. Perbedaan estimasi dan klastering
estimasi digunakan untuk menerka sebuah nilai yang belum di ketahui. Sedangkan klastering merupakan pengelompokan data berdasarkan kemiripan objek data.

10. Perbedaan Supervised dan unsupervised
Supervised dapat mengenali data dari label khusus yang telah di berikan sebelumnya. Sedangkan unsupervised adalah mengenali data secara real-time begitu data di sajikan.

11. Tahapan utama proses data mining

- Masukkan (dataset)
- Metode (Algoritma data mining)
- Keluaran / pola / model / pengetahuan
- Evaluasi (akurasi, error, dan lain lain.