

Nama: Muhammad Fikri Ridwan  
Nim : A11.2019.11976  
Data Mining – 4611

## Tugas

1. Sebutkan 5 peran utama data mining!
2. Algoritma apa saja yang dapat digunakan untuk 5 peran utama data mining di atas?
3. Jelaskan perbedaan estimasi dan prediksi!
4. Jelaskan perbedaan prediksi dan klasifikasi!
5. Jelaskan perbedaan klasifikasi dan klustering!
6. Jelaskan perbedaan klustering dan association!
7. Jelaskan perbedaan supervised dan unsupervised learning!
8. Sebutkan tahapan utama proses data mining!

Jawab

1. Estimation  
Prediction  
Classification  
Clustering  
Association
2. Algoritma Data Mining :
  - Estimation : Linier Regression, Neural Network, Support Vector Machine, dll
  - Prediction : Linier Regression, Neural Network, Support Vector Machine, dll
  - Classification : Naive Bayes, K-Nearest Neighbor, C4.5, ID3, CART, Linear Discriminant Analysis, dll
  - Clustering : K-Means, K-Medoids, Self-Organizing Map(SOM), Fuzzy C-Means, dll
  - Association : FP-Growth, A Priori, dll
3. *Estimation* (estimasi), untuk menerka sebuah nilai yang belum diketahui, misal menerka penghasilan seseorang ketika informasi mengenai orang tersebut diketahui. Metode yang digunakan antara lain *Point Estimation* dan *Confidence Interval Estimations*, *Simple Linear Regression* dan *Correlation*, dan *Multiple Regression*.  
  
*Prediction* (prediksi), untuk memperkirakan nilai masa mendatang, misal memprediksi stok barang satu tahun ke depan. Fungsi ini mencakup metode *Neural Network*, *Decision Tree*, dan *k-Nearest Neighbor*.
4. *Prediction* (prediksi), untuk memperkirakan nilai masa mendatang, misal memprediksi stok barang satu tahun ke depan. Fungsi ini mencakup metode *Neural Network*, *Decision Tree*, dan *k-Nearest Neighbor*.

*Classification* (klasifikasi), merupakan proses penemuan model atau fungsi yang menjelaskan atau membedakan konsep atau kelas data, dengan tujuan untuk dapat memperkirakan kelas dari suatu objek yang labelnya tidak diketahui. Metode yang digunakan antara lain *Neural Network*, *Decision Tree*, *k-Nearest Neighbor*, dan *Naive Bayes*.

5. *Classification* (klasifikasi), merupakan proses penemuan model atau fungsi yang menjelaskan atau membedakan konsep atau kelas data, dengan tujuan untuk dapat memperkirakan kelas dari suatu objek yang labelnya tidak diketahui. Metode yang digunakan antara lain *Neural Network*, *Decision Tree*, *k-Nearest Neighbor*, dan *Naive Bayes*.

*Clustering* (pengelompokan), yaitu pengelompokan mengidentifikasi data yang memiliki karakteristik tertentu. Metode dalam fungsi ini diantaranya *Hierarchical Clustering*, metode *K-Means*, dan *Self Organizing Map* (SOM)

6. *Clustering* (pengelompokan), yaitu pengelompokan mengidentifikasi data yang memiliki karakteristik tertentu. Metode dalam fungsi ini diantaranya *Hierarchical Clustering*, metode *K-Means*, dan *Self Organizing Map* (SOM)

*Association* (asosiasi), dinamakan juga analisis keranjang pasar dimana fungsi ini mengidentifikasi item-item produk yang kemungkinan dibeli konsumen bersamaan dengan produk lain. Metode atau algoritma dalam fungsi ini adalah *Apriori*, *Generalized Sequential Pattern (GSP)*, *FP-Growth* dan *GRI algorithm*

7. Pendekatan supervised learning mempunyai input dan output yang dapat dibuat menjadi suatu model hubungan matematis sehingga mampu melakukan prediksi dan klasifikasi berdasarkan data yang telah ada sebelumnya.

Pendekatan unsupervised learning tidak menggunakan data latih atau data training untuk melakukan prediksi maupun klasifikasi. Berdasarkan model matematisnya, algoritma ini tidak memiliki target variabel. Salah satu tujuan dari algoritma ini adalah mengelompokkan objek yang hampir sama dalam suatu area tertentu.

8. Pembersihan data  
Integrasi data  
Transformasi data  
Aplikasi teknik data mining  
Evaluasi pola yang ditemukan  
Presentasi pola yang ditemukan untuk menghasilkan aksi