

Nama : Wisnu Putra Pratama

NIM : A11.2018.11267

1. Sebutkan 5 peran utama data mining!

- Estimasi
- Prediksi
- Klasifikasi
- Klastering
- Asosiasi

2. Algoritma apa saja yang dapat digunakan untuk 5 peran utama data mining di atas?

- Estimasi : Linier Regression, Neural Network, Support Vector Machine, dll
- Prediksi : Linier Regression, Neural Network, Support Vector Machine, dll
- Klasifikasi : Naive Bayes, K-Nearest Neighbor, C4.5, ID3, CART, Linear Discriminant Analysis, dll
- Klastering : K-Means, K-Medoids, Self-Organizing, Map(SOM), Fuzzy C-Means, dll
- Asosiasi : FP-Growth, A Priori, dll

3. Jelaskan perbedaan estimasi dan prediksi!

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.

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. Jelaskan perbedaan Estimasi dan klasifikasi

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.

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. Jelaskan perbedaan klasifikasi dan klastering!

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.

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

6. Jelaskan perbedaan klastering dan prediksi!

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

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.

7. Jelaskan perbedaan supervised dan unsupervised learning!

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. Sebutkan tahapan utama proses data mining!

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