

jawab

1. Data mining, secara sederhana, adalah istilah yang digunakan untuk menjelaskan proses pencarian atau penambangan knowledge dari data yang sangat besar di suatu Perusahaan atau lembaga untuk menentukan pola perilaku konsumen atau pegawai

2. Apa saja proses tahapan data mining?

- Pembersihan data (untuk membuang data yang tidak konsisten dan noise)
- Integrasi data (penggabungan data dari beberapa sumber)
- Transformasi data (data diubah menjadi bentuk yang sesuai untuk di-mining)
- Aplikasi teknik Data Mining, proses ekstraksi pola dari data yang ada
- Evaluasi pola yang ditemukan (proses interpretasi pola menjadi pengetahuan yang dapat digunakan untuk mendukung pengambilan keputusan)
- Presentasi pengetahuan (dengan teknik visualisasi)

3. Description digunakan untuk mengidentifikasi pola yang sering muncul dan mengubah pola tersebut menjadi aturan yang dapat digunakan untuk memudahkan suatu aktivitas.

Classification digunakan jika atributnya bisa berupa nominal atau numerik tetapi labelnya harus berupa nominal. Pada klasifikasi dilakukan pengelompokkan berdasarkan hubungan antara variabel kriteria dengan variabel target

Prediction digunakan jika atributnya numerik dan labelnya numerik serta terdapat time series pada atributnya. Peran data mining prediksi ini secara umum hampir sama dengan klasifikasi.

Estimation digunakan jika dataset atributnya numerik dan labelnya numerik.

Clustering digunakan jika dataset tidak memiliki label dan atributnya numerik.

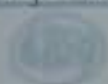
Pengklasteran merupakan pengelompokkan data yang memiliki kemiripan nilai

Association digunakan jika dataset ingin diketahui seberapa dekat hubungan antar atributnya

4. Estimation (estimasi), untuk menerka sebuah nilai yang belum diketahui, misal menerka penghasilan seseorang ketika informasi mengenai orang tersebut diketahui.
2. Forecasting (prediksi), untuk memperkirakan nilai masa mendatang, misal memprediksi stok barang satu tahun ke depan.

5. Forecasting Adalah Suatu Metode Dimana Kita Dapat Memperkirakan Nilai Dari Suatu Populasi Dengan Menggunakan Nilai Dari Sambil. Sedangkan, Klasifikasi Adalah Metode Penganalisaan Data Yang Sering Dimasukkan Sebagai Salah Satu Metode Data Mining.

6. Klasifikasi adalah pemrosesan untuk menemukan sebuah model atau fungsi yang menjelaskan dan mencirikan konsep atau kelas data, untuk kepentingan tertentu. sedangkan clustering digunakan untuk pengelompokkan data berdasarkan kemiripan pada objek data dan sebaliknya meminimalkan kemiripan terhadap kluster yang lain



7. Klustering Proses pengumpulan data dengan jumlah besar sedangkan association teknik data mining untuk menemukan aturan asosiasi antara suatu kombinasi item
8. estimasi mirip dengan algoritma klasifikasi, tapi variabel target adalah berupa bilangan numerik (kontinyu) dan bukan kategorikal (nominal atau diskrit) sedangkan algoritma klasifikasi menggunakan target/class/label berupa nilai kategorikal.
9. Estimasi Adalah Suatu Metode Dimana Kita Dapat Memperkirakan Nilai Dari Suatu Populasi Dengan Menggunakan Nilai Dari Sambilan. Sedangkan Klustering Adalah Metod Penganalisaan Data Yang Sering Dimasukkan Sebagai Salah Satu Metode Data Mining.
10. Supervised mengenali data dari label khusus yang telah diberikan sebelumnya, sedangkan unsupervised mengenali data secara real-time begitu data disajikan
11. Pembersihan data (untuk membuang data yang tidak konsisten dan noise)
- Integrasi data (penggabungan data dari beberapa sumber)
- Transformasi data (data diubah menjadi bentuk yang sesuai untuk di-mining)
- Aplikasi teknik Data Mining, proses ekstraksi pola dari data yang ada

