

Tugas Sains Data Kelas C

Nama: Rizka Iyiba

Nim: 19.01.013.116

1) Data mining adalah sebuah proses ilmu yang melakukan ekstraksi dengan cara penggalian informasi, mengidentifikasi pola, melakukan pemakaian data historis untuk mendapatkan informasi penting dari sebuah data yang sangat besar dan berguna juga untuk pengambilan keputusan suatu perusahaan & organisasi.

2) Konsep alur proses data mining:

a) Himpunan Data

b) Metode Data mining

c) Pengetahuan

Beberapa proses yang lain juga yakni:

a) Deskripsi (mengidentifikasi pola)

b) Prediksi (Mengklasifikasi)

c) Estimasi (variabelnya lebih ke arah numerik)

d) Klasifikasi (model fungsi)

e) Clustering (Pengelompokan data)

f) Asosiasi (menemukan atribut yg muncul dlm waktu)

3) 5 peran utama data mining:

a. Estimasi

b. Forecasting

c. Klasifikasi

d. Clustering

e. Asosiasi

4) Perbedaan estimasi dan forecasting

estimasi itu untuk mencari sebuah nilai yang belum diketahui sebelumnya. ~~Itu~~ ataupun suatu penghitungan parameter / jumlah populasi yang belum diketahui. Sedangkan forecasting itu ialah sebuah proses untuk memperkirakan nilai masa mendatang atau memperkirakan nilai untuk variabel dependen yang didasarkan pada beberapa nilai yang tidak teramati dari variabel yang sifatnya independen.

5) Perbedaan Forecasting dan klasifikasi

Forecasting itu metode untuk memperkirakan nilai dari suatu populasi menggunakan nilai dari sampel. Sedangkan klasifikasi itu metode

menganalisa data yang sering dimasukkan dalam metode data mining

⑥ Perbedaan klasifikasi dan klastering!

Klasifikasi digunakan untuk menentukan model untuk kepentingan tertentu sedangkan klastering merupakan teknik untuk mengelompokkan data berdasarkan kemiripan.

⑦ Perbedaan klastering dan asosiasi!

Klastering dipakai saat tidak diketahuinya bagaimana data harus dikelompokkan sehingga menghasilkan data yang sudah dikelompokkan. Sedangkan klasifikasi terdapat informasi mengenai bagaimana data tersebut dikelompokkan ~~yang~~ untuk menghasilkan ~~klasifikasi~~ hubungan yang menanak di antara berbagai variabel dalam database yang sangat besar

⑧ Perbedaan estimasi dan klasifikasi!

Estimasi ialah metode dimana kita dapat memperkirakan nilai dari suatu populasi dengan menggunakan nilai dari sampel. Sedangkan klasifikasi dan estimasi hampir mirip, tetapi variabel targetnya berupa bilangan numerik (kontinu) dan bukan kategorikal (nominal / diskrit)

⑨ Perbedaan estimasi dan klastering!

Estimasi metode untuk menyelesaikan yang berkaitan dengan memperkirakan seberapa banyak hasil produksi sedangkan klastering merupakan salah satu aktifitas analisis data untuk mengelompokkan data ke dalam beberapa kategori / kluster.

⑩ Perbedaan supervised dan unsupervised learning!

Supervised learning ialah sebuah pendekatan yang sudah terdapat data training dan terdapat variable yang ditargetkan untuk mengelompokkan suatu data ke data yang sudah ada.

Sedangkan, untuk unsupervised learning tidak memiliki data training, sehingga dari data yang ada akan dikelompokkan menjadi 2 bagian atau 3 bagian data dan seterusnya.

⑪ Tahapan Utama Proses data mining

Pertama yaitu pembersihan data, lalu integrasi data, transformasi data, lalu proses ekstraksi pola, evaluasi pola, dan presentasi pengetahuan