

Nama : Ismayana Dwi Kartini

NIM : 211002019

Prodi : Teknik Metalurgi

### Tugas

1. Jelaskan dengan kalimat sendiri apa yang dimaksud dengan data mining?

= Definisi data mining, secara sederhana, adalah istilah yang digunakan untuk menjelaskan proses pencarian atau penambangan knowledge dari data yang sangat besar. Secara teknis, data mining adalah proses yang memanfaatkan teknik-teknik statistik, matematika, dan kecerdasan buatan untuk mengekstrak dan mengidentifikasi informasi knowledge selanjutnya (atau pola-pola) yang berasal dari sekumpulan data yang sangat besar. Data mining adalah suatu proses penambangan informasi penting dari suatu data.

2. Sebutkan konsep alur proses data mining!

= Konsep alur proses data mining memiliki tahap sebagai berikut:

a.) Pembersihan data

b.) Integrasi data

c.) Transformasi data



3. Sebutkan 5 peran utama data mining!

= Peran utama data mining sebagai berikut:

a.) Description

b.) Classification

c.) Prediction

d.) Estimation

e.) Clustering

4. Jelaskan perbedaan estimasi dan forecasting!

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

5. Jelaskan perbedaan forecasting dan klasifikasi!

= Forecasting adalah suatu metode dimana kita dapat memperkirakan nilai dari suatu populasi dengan menggunakan nilai dari sampel. Sedangkan, klasifikasi adalah metode penganalisaan data yang sering dimasukkan sebagai salah satu metode data mining.

6. Jelaskan perbedaan klasifikasi dan klastering!

= Klasifikasi adalah pemrosesan untuk menemukan sebuah model atau fungsi yang menjelaskan dan



☐ menceritakan konsep atau kelas data untuk kepentingan tertentu. Sedangkan clustering digunakan untuk pengelompokan data berdasarkan kemiripan pada objek data dan sebaliknya meminimalkan kemiripan terhadap kluster yang lain.

7. Jelaskan perbedaan clustering dan association!

☐ = Clustering dipakai ketika tidak diketahuinya bagaimana data harus dikelompokkan. Jumlah kelompok ~~dian~~ diasumsikan sendiri tanpa ditentukan terlebih dahulu. Keluaran pendekatan ini adalah data yang sudah dikelompokkan. Sedangkan classification terdapat informasi mengenai bagaimana data tersebut dikelompokkan. Kemudian dilakukan training pada sistem dengan data yang sudah diberikan label (ke dalam kelompok manakah data tersebut dikelompokkan), selanjutnya sistem akan mengklasifikasikan data-data yang baru ke dalam kelompok yang ada.

8. Jelaskan perbedaan estimasi dan klasifikasi!

☐ = Algoritma estimasi mirip dengan algoritma klasifikasi, tapi variabel target adalah berupa bilangan numerik (kontinu) dan bukan kategorikal (nominal atau diskrit) sedangkan algoritma klasifikasi menggunakan target/class/label berupa nilai kategorikal (nominal).



9. Jelaskan perbedaan estimasi dan klastering!

= Estimasi adalah suatu metode dimana kita dapat memperkirakan nilai dari suatu populasi dengan menggunakan nilai dari sampel. Sedangkan klastering adalah metode pengorganisasian data yang sering dimasukkan sebagai salah satu metode data mining.

10. Jelaskan perbedaan supervised dan unsupervised learning!

= Baik supervised maupun unsupervised learning adalah pendekatan yang dilakukan algoritma komputer dalam mengenali pola pada data. Supervised mengenali data dari label khusus yang telah diberikan sebelumnya, sedangkan unsupervised mengenali data secara real-time begitu data disajikan.

11. Sebutkan tahapan utama proses data mining!

= Tahapan utama proses data mining, yaitu sebagai berikut:

a.) Pembersihan data

b.) Integrasi data

c.) Transformasi data

d.) Aplikasi teknik data mining proses ekstraksi pola dari data yang ada.