

Nama : Ferza Reyaldi  
NIM : 09021281924060  
Mata Kuliah : Data Mining

### **Tugas 1**

1. Sebutkan 5 peran utama data mining!

**Jawab:**

Peran utama data mining menurut Buulolo (2020), yaitu:

- 1) Description;
- 2) Classification;
- 3) Prediction;
- 4) Estimation;
- 5) Clustering;
- 6) Association.

2. Jelaskan perbedaan estimasi dan prediksi!

**Jawab:**

Estimasi adalah salah satu metode data mining yang digunakan untuk mengetahui atau menemukan nilai yang belum diketahui dengan cara mengidentifikasi dan menganalisis hubungan antar variabel.

Sedangkan,

Prediksi adalah salah satu metode data mining yang digunakan menganalisis peristiwa masa lalu atau contoh dalam urutan yang tepat untuk memprediksi peristiwa masa depan.

3. Jelaskan perbedaan prediksi dan klasifikasi!

**Jawab:**

Prediksi adalah salah satu metode data mining yang digunakan menganalisis peristiwa masa lalu atau contoh dalam urutan yang tepat untuk memprediksi peristiwa masa depan.

Sedangkan,

Klasifikasi adalah salah satu metode data mining yang digunakan untuk mengelompokkan data berdasarkan kelas-kelas, kemudian mempelajari atribut-atribut ada, dan dihasilkan rules berdasarkan hal tersebut yang digunakan untuk mengklasifikasikan data yang baru.

4. Jelaskan perbedaan klasifikasi dan klastering!

**Jawab:**

Klasifikasi adalah salah satu metode data mining yang digunakan untuk mengelompokkan data berdasarkan kelas-kelas, kemudian mempelajari atribut-atribut ada, dan dihasilkan rules berdasarkan hal tersebut yang digunakan untuk mengklasifikasikan data yang baru.

Sedangkan,

Klastering adalah salah satu metode data mining yang digunakan untuk mengelompokkan data berdasarkan kemiripan atribut-atributnya.

5. Jelaskan perbedaan klustering dan association!

**Jawab:**

Klustering adalah salah satu metode data mining yang digunakan untuk mengelompokkan data berdasarkan kemiripan atribut-atributnya.

Sedangkan,

Association adalah salah satu metode data mining yang membantu untuk menemukan hubungan antara dua atau lebih item. Ia menemukan pola tersembunyi dalam kumpulan data.

6. Jelaskan perbedaan estimasi dan klasifikasi!

**Jawab:**

Estimasi adalah salah satu metode data mining yang digunakan untuk mengetahui atau menemukan nilai yang belum diketahui dengan cara mengidentifikasi dan menganalisis hubungan antar variabel.

Sedangkan,

Klasifikasi adalah salah satu metode data mining yang digunakan untuk mengelompokkan data berdasarkan kelas-kelas, kemudian mempelajari atribut-atribut ada, dan dihasilkan rules berdasarkan hal tersebut yang digunakan untuk mengklasifikasikan data yang baru.

7. Jelaskan perbedaan estimasi dan klustering!

**Jawab:**

Estimasi adalah salah satu metode data mining yang digunakan untuk mengetahui atau menemukan nilai yang belum diketahui dengan cara mengidentifikasi dan menganalisis hubungan antar variabel.

Sedangkan,

Klustering adalah salah satu metode data mining yang digunakan untuk mengelompokkan data berdasarkan kemiripan atribut-atributnya.

8. Jelaskan perbedaan supervised dan unsupervised learning!

**Jawab:**

Perbedaan utama antara kedua pendekatan tersebut adalah penggunaan dataset berlabel. Sederhananya, supervised learning menggunakan data input dan output berlabel, sedangkan unsupervised learning menggunakan data yang tidak berlabel.

Jika ditinjau dari sisi kompleksitas, supervised learning adalah metode sederhana untuk pembelajaran mesin. Sedangkan untuk unsupervised learning, diperlukan alat canggih untuk bekerja dengan sejumlah besar data yang tidak diklasifikasikan. Model unsupervised learning rumit secara komputasi karena mereka membutuhkan pelatihan yang besar untuk menghasilkan hasil yang diinginkan.

9. Sebutkan tahapan utama proses data mining!

**Jawab:**

Tahapan utama proses data mining (berdasarkan CRISP-DM), yaitu:

- 1) Business understanding;
- 2) Data understanding;

- 3) Data preparation;
- 4) Modelling;
- 5) Evaluation;
- 6) Deployment.