Nama : Genta Andrean Ardinata

Nim : A11.2022.14092

Klp : A11.4509

Tugas 4

1. **Sebutkan 5 peran utama data mining!**

1.Estimasi

2. Prediksi

3. Klasifikasi

4. Klastering

5. Asosiasi

1. **algoritma apa saja yang dapat digunakan untuk 5 peran utama data mining di atas?**

1. Estimation (Estimasi): Linear Regression, Neural Network, Support Vector Machine, etc

2. Prediction/Forecasting (Prediksi/Peramalan): Linear Regression, Neural Network, Support Vector Machine, etc

3. Classification (Klasifikasi): Naive Bayes, K-Nearest Neighbor, C4.5, ID3, CART, Linear Discriminant Analysis, etc

4. Clustering (Klastering): K-Means, K-Medoids, Self-Organizing Map (SOM), Fuzzy C-Means, etc

5. Association (Asosiasi): FP-Growth, A Priori, etc

1. **Jelaskan perbedaan estimasi dan prediksi!**

Algoritma prediksi/forecasting sama dengan algoritma estimasi di mana label/target/class bertipe numerik, bedanya adalah data yang digunakan merupakan data rentet waktu (data time series).

1. **Jelaskan perbedaan estimasi dan klasifikasi!**

Algoritma estimasi mirip dengan algoritma klasifikasi, bedanya variabel target adalah berupa bilangan numerik (kontinyu) dan bukan kategorikal (nominal atau diskrit).

1. **Jelaskan perbedaan klasifikasi dan klastering!**

Klasifikasi adalah algoritma yang menggunakan data dengan target/class/label berupa nilai kategorikal (nominal), Perbedaan utama algoritma klastering dengan klasifikasi adalah klastering tidak memiliki target/class/label, jadi termasuk unsupervised learning.

1. **Jelaskan perbedaan klastering dan prediksi!**

Klastering adalah pengelompokkan data, hasil observasi dan kasus ke dalam class yang mirip, sedangkan prediksi kadang digunakan juga untuk klasifikasi, tidak hanya untuk prediksi time series, karena sifatnya yang bisa menghasilkan class berdasarkan berbagai atribut yang kita sediakan.

1. **Jelaskan perbedaan supervised dan unsupervised learning!**

supervised learning memiliki Variabel yang menjadi target/label/class ditentukan, sedangkan unsupervised learning Variable (atribut) yang menjadi target/label/class tidak ditentukan (tidak ada).

1. **Sebutkan tahapan utama proses data mining!**

* INPUT
* METODE
* OUTPUT
* EVALUATION