

Nama : Muhammad Alif al Husain

NIM : A11.2022.14718

Tugas4

- Soal
 1. sebutkan 5 peran Utama data mining
 2. algoritma apa saja yang dapat digunakan untuk 5 peran Utama data mining
 3. jelaskan perbedaan estimasi dan prediksi
 4. perbedaan estimasi dan klasifikasi
 5. perbedaan klasifikasi dan klustering
 6. perbedaan klustering dan prediksi
 7. perbedaan supervised dan unsupervised learning
 8. sebutkan tahapan Utama proses data mining.
- Jawab
 1. 5 peran Utama dari data mining adalah estimasi, prediksi, klasifikasi, klustering, dan asosiasi.
 - 2.
- Estimasi

Algoritma populer :

 - Regresi Linear
Untuk memperkirakan hubungan antara variabel dependen dan independen.
 - Regresi Lasso
Regresi dengan regularisasi untuk mengatasi overfitting.
 - Neural Networks
Untuk estimasi kompleks seperti time series.
- Prediksi

Algoritma populer :

 - Regresi Logistik
Prediksi nilai dan probabilitas.
 - Random Forest
Model ansamabel untuk prediksi
 - ARIMA
Prediksi deret waktu.
- Klasifikasi

Algoritma populer :

 - K-NN
Mengklasifikasikan data berdasarkan tetangga terdekat.
 - Naïve Bayes
Klasifikasi berbasis probabilitas.
 - Random Forest
Ansambel beberapa pohon keputusan.
- Klustering

Algoritma populer :

 - K-Means
Memisahkan data ke dalam kelompok berdasarkan centroid.
 - Hierarchical Clustering
Membangun dendrogram untuk representasi kluster.

- GMM
Kluster berdasarkan distribusi probabilitas.
- Asosiasi
 - Algoritma populer :
 - Apriori
Mencari aturan asosiasi berdasarkan frekuensi.
 - ECLAT
Pendekatan berbasis set item untuk mencari asosiasi.
- 3. Algoritma prediksi/forecasting mirip dengan algoritma estimasi dimana label/target bertipe numerik akan tetapi memiliki perbedaan data yang digunakan merupakan data rentet Waktu. Semua algoritma estimasi dapat digunakan untuk prediksi.
- 4. Algoritma estimasi mirip dengan algoritma klasifikasi, akan tetapi variable target adalah berupa bilangan numerik dan bukan kategorikal.
- 5. Perbedaan klasifikasi dan klustering adalah jika klasifikasi memiliki tujuan untuk memprediksi kelas data baru sedangkan klustering mengelompokkan data serupa.
- 6. Perbedaan klustering dan prediksi adalah jika klustering mengelompokkan data serupa sedangkan prediksi memprediksi nilai atau hasil.
- 7. Supervised Learning (Pembelajaran dengan guru)
Sebagian besar algoritma yang ada didalam data mining adalah supervised learning sebagai contoh : estimation, prediction/forecasting, classification.
Algoritma melakukan proses belajar berdasarkan nilai dari variable target yang terasosiasi engan nilai variable predictor.
- Unsupervised Learning(Pembelajaran tanpa guru)
Algoritma data mining yang mencari pola dari semua variabel(attribute)
Algoritma clustering adalah contoh dari algoritma unsupervised learning.
- 8. Tahapan utama Data Mining
 - Pengumpulan data
 - Pembersihan data
 - Transformasi data
 - Pemilihan data
 - Pemodelan
 - Evaluasi model
 - Visualisasi
 - Deployment dan monitoring