Nama : Rahmat Suci Setyanto

NIM : A11.2022.14710

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 klastering
  6. perbedaan klastering 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 kompelks 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 dendrogam untuk representasi kluster.

* + Asosiasi
* GMM

Kluster berdasarkan distribusi probabilitas.

Algoritma populer :

* Apriori

Mencari aturan asosiasi berdasarkan frekuensi.

* ECLAT

Pendekatan berbasis set item untuk mencari asosiasi.

1. Algoritma prediksi/forecasting mirip dengan algoritma estimasi dimana label/target

bertipe numerik akan tertapi memiliki perbedaan data yang digunakan merupakan data rentet Waktu. Semua algoritma estimasi dapat digunakan untuk prediksi.

1. Algoritma estimasi mirip dengan algoritma klasifikasi,akan tetapi variable target adalah berupa bilangan numerik dan bukan kategorikal.
2. Perbedaan klasifikasi dan klastering adalah jika klasifikasi memiliki tujuan untuk memprediksi kelas data baru sedangkan klastering mengelompokkan data serupa.
3. Perbedaan klastering dan prediksi adalah jika klastering mengelompokkan data serupa sedangkan prediksi memprediksi nilai atau hasil.
4. Supervised Learning (Pembelajaran dengan guru)

Sebagain 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 variabkle(atribute) Algoritma clustering adalah contoh dari algoritma unsupervised learning.

1. Tahapan utama Data Mining
   * Pengumpulan data
   * Pembersihan data
   * Transformasi data
   * Pemilihan data
   * Pemodelan
   * Evaluasi model
   * Visualisasi
   * Deployment dan monitoring