

Nama : Widiya Lis Susana
NIM : 211001002

SAINS DATA

Data Mining adalah disiplin ilmu yang mempelajari metode untuk mengekstrak pengetahuan atau menemukan pola dari suatu data yang besar.

konsep Data Mining :

- Seleksi
- Preprocessing
- Transformasi
- Data Mining
- Interpretasi dan Evaluasi.

3. 5 Peran utama data mining :

- Classification
- Clustering
- Association
- Sequencing
- Forecasting / Prediksi.

4. Perbedaan Estimasi dan Prediksi

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

5. Perbedaan Forecasting dan Klasifikasi

- klasifikasi, memprediksi label yang bersifat kategori (diskrit dan tidak berurutan).
- Prediksi, memodelkan fungsi nilai yang kontinyu.

6. Perbedaan klasifikasi dan klastering

- klasifikasi adalah pemrosesan untuk menemukan sebuah model atau fungsi yang menjelaskan dan mencirikan konsep atau kelas data untuk kepentingan tertentu. sedangkan
- klastering digunakan untuk pengelompokan data berdasarkan kemiripan pada objek data dan sebaliknya meminimalkan kemiripan terhadap kluster yang lain.

7. Perbedaan Klastering dan Association

- Klastering adalah pengelompokan data tanpa berdasarkan kelas data tertentu ke dalam kelas objek yang sama.
- Association adalah menemukan atribut yang muncul dalam suatu waktu.

8. Perbedaan Estimasi dan Klasifikasi

- Estimasi mirip dengan algoritma klasifikasi tapi variabel target adalah berupa bilangan numerik (kontinu) dan bukan kategorikal (nominal atau diskrit).
- Klasifikasi menggunakan target / class / label berupa nilai kategorikal (nominal).

9. Perbedaan Estimasi dan Klastering

- Estimasi adalah suatu metode dimana kita dapat memperkirakan nilai dari suatu populasi dengan menggunakan nilai dari sampel.
- Klastering adalah metode penganalisaan data yang sering dimasukkan sebagai salah satu metode data mining.

10. Perbedaan supervised dan unsupervised learning

- Supervised, mengenali data dari label khusus yang telah diberikan sebelumnya.
- Unsupervised mengenali data secara real-time begitu data disajikan.

11.