

Modul Praktikum Data Mining

A. Tujuan Pembelajaran

- Mahasiswa mampu memahami konsep dasar data mining.
- Mahasiswa dapat menginstal dan menggunakan perangkat lunak data mining.
- Mahasiswa mampu melakukan proses dasar data mining (klasifikasi, clustering, asosiasi).
- Mahasiswa dapat menganalisis hasil keluaran dari proses data mining.

B. Dasar Teori

Data Mining adalah proses mengekstrak pengetahuan dari kumpulan data besar menggunakan teknik statistik, machine learning, dan database.

Tahapan Data Mining (CRISP-DM):

- • Business Understanding
- • Data Understanding
- • Data Preparation
- • Modeling
- • Evaluation
- • Deployment

Metode umum dalam data mining meliputi:

- • Klasifikasi (Decision Tree, Naïve Bayes, KNN)
- • Clustering (K-Means, Hierarchical)
- • Asosiasi (Apriori, FP-Growth)

C. Alat dan Bahan

- Laptop/PC dengan RAM minimal 4GB.
- Software: Python (Anaconda, Jupyter Notebook) / Weka / RapidMiner.
- Dataset (contoh: Iris Dataset, Dataset Penjualan, atau Dataset Customer).

D. Langkah Praktikum

Praktikum 1: Klasifikasi dengan Decision Tree (Dataset Iris)

- Buka RapidMiner / Python Jupyter Notebook.
- Import dataset Iris.csv.
- Lakukan preprocessing (cek missing value, normalisasi jika perlu).

- Pilih metode Decision Tree untuk klasifikasi.
- Bagi data menjadi training (70%) dan testing (30%).
- Jalankan proses mining.
- Catat hasil akurasi model.

Praktikum 2: Clustering dengan K-Means (Dataset Customer Segmentation)

- Import dataset Customer.csv.
- Pilih metode K-Means Clustering.
- Tentukan jumlah cluster (misal: $k=3$).
- Jalankan proses clustering.
- Visualisasikan hasil clustering.
- Analisis hasil cluster.

Praktikum 3: Association Rule (Apriori) (Dataset Transaksi Penjualan)

- Import dataset Transaksi Penjualan (format CSV).
- Pilih metode Apriori Algorithm.
- Atur parameter min support (0.2) dan min confidence (0.6).
- Jalankan proses mining.
- Catat aturan asosiasi yang muncul.

E. Tugas/LKPD

- Lakukan klasifikasi dataset Iris dengan metode Naïve Bayes. Bandingkan hasil akurasinya dengan Decision Tree.
- Gunakan dataset pelanggan untuk clustering dengan jumlah cluster berbeda ($k=2$, $k=4$). Analisis perbedaan hasilnya.
- Gunakan dataset transaksi, ubah nilai minimum support & confidence, kemudian analisis perubahan aturan asosiasi.

F. Evaluasi

- Apa perbedaan utama antara klasifikasi dan clustering?
- Sebutkan contoh penerapan nyata dari algoritma Apriori dalam bisnis retail.
- Mengapa preprocessing data penting dalam data mining?