**LAPORAN PRAKTIKUM**

**DATA MINING**

**Disusun Oleh :**

Nama : Rifky Syah Fredian

NIM : 231011403277

Kelas : 04TPLP027

**TEKNIK INFORMATIKA**

**FAKULTAS TEKNIK**

**UNIVERSITAS PAMULANG**

JL. RAYA PUSPITEK, BUARAN, KEC. PAMULANG, KOTA TANGGERANG SELATAN

BANTEN.

LAPORAN PRAKTIKUM DATA MINING

Nama: Rifky Syah Fredian

NIM: 231011403277

Kelas: 04TPLP027

Mata Kuliah: Data Mining

# Judul: Analisis Lalu Lintas Jaringan Menggunakan Dataset "Benign Traffic"

## 1. Pendahuluan

Data mining merupakan proses penting dalam mengekstraksi informasi tersembunyi dari data dalam jumlah besar. Dalam konteks keamanan jaringan, data mining dapat membantu dalam mendeteksi pola-pola trafik yang mencurigakan atau berbahaya. Salah satu metode yang digunakan adalah analisis lalu lintas jaringan menggunakan dataset yang mencerminkan aktivitas jaringan. Pada praktikum ini, digunakan dataset "Benign Traffic" yang berisi data lalu lintas jaringan yang dianggap tidak berbahaya. Tujuan dari praktikum ini adalah untuk memahami cara membaca, menampilkan, dan melakukan evaluasi dasar terhadap data jaringan menggunakan Python dan beberapa pustaka pendukung seperti pandas, seaborn, dan scikit-learn.

## 2. Langkah-langkah Praktikum

a. Import Library

Langkah pertama adalah mengimpor pustaka yang dibutuhkan untuk proses analisis. Pustaka yang digunakan antara lain pandas untuk manipulasi data, seaborn dan matplotlib untuk visualisasi, serta scikit-learn untuk evaluasi model. Namun, saat pertama kali dijalankan, beberapa pustaka seperti matplotlib dan seaborn belum terinstal sehingga menimbulkan error.

b. Membaca Dataset

Dataset dibaca menggunakan fungsi pd.read\_csv("Benign Traffic.csv"). Dataset ini memuat berbagai informasi mengenai lalu lintas jaringan, termasuk IP sumber, IP tujuan, durasi koneksi, jumlah paket, dan lain-lain.

c. Menampilkan Data

Setelah data berhasil dimuat ke dalam DataFrame, langkah selanjutnya adalah menampilkan isi dari dataset menggunakan perintah `dataset` untuk mengetahui struktur kolom, jumlah data, serta jenis data yang tersedia.

d. Visualisasi dan Evaluasi Data

Setelah memahami isi data, dilakukan evaluasi awal menggunakan confusion matrix dan classification report dari library scikit-learn. Selain itu, digunakan heatmap dari seaborn untuk memvisualisasikan performa klasifikasi jika data label tersedia. Namun karena adanya error module yang belum terinstal, proses ini belum berjalan dengan sempurna hingga semua dependensi berhasil diinstal.

## 3. Kendala yang Dihadapi

Beberapa kendala yang ditemui saat praktikum antara lain:  
- Error "ModuleNotFoundError" muncul karena pustaka seperti matplotlib dan seaborn belum terinstal.  
- Beberapa data memerlukan proses pra-pemrosesan lebih lanjut agar dapat digunakan dalam pelatihan model.  
- Tidak semua data memiliki label yang jelas, sehingga menyulitkan dalam melakukan evaluasi klasifikasi secara menyeluruh.

## 4. Solusi dan Tindakan

Untuk mengatasi error import pustaka, dilakukan instalasi menggunakan perintah:  
- `pip install matplotlib`  
- `pip install seaborn`  
  
Selain itu, perlu dilakukan analisis lebih lanjut pada dataset untuk melakukan normalisasi data, mengisi nilai yang hilang, serta memetakan label jika ingin melakukan klasifikasi dengan algoritma supervised.

## 5. Kesimpulan

Praktikum ini memberikan pemahaman dasar tentang cara membaca dan mengevaluasi dataset menggunakan Python. Walaupun terdapat kendala teknis pada instalasi pustaka, secara keseluruhan proses analisis berjalan dengan baik. Ke depannya, evaluasi lanjutan seperti training model machine learning dapat dilakukan untuk deteksi lalu lintas jaringan berbahaya.