

# **LAPORAN PRAKTIKUM**

## **DATA MINING**



Disusun Oleh:

Nama : Zidan Faturrahman

NIM : 231011401114

Kelas : 04TPLP027

**UNIVERSITAS PAMULANG**

FAKULTAS ILMU KOMPUTER

PROGRAM STUDI TEKNIK INFORMATIKA

2025

# **BAB I**

## **PENDAHULUAN**

Data mining adalah proses menemukan pola menarik dan pengetahuan tersembunyi dari kumpulan data yang besar. Dalam praktik ini, data mining digunakan untuk menganalisis lalu lintas jaringan dan mendeteksi adanya aktivitas berbahaya, seperti serangan DDoS (Distributed Denial of Service). Laporan ini berfokus pada analisis dua dataset, yaitu data benign (lalu lintas normal) dan data serangan DDoS ICMP Flood.

## **BAB II**

### **DASAR TEORI**

#### **1. Data Mining**

Data mining merupakan bagian dari ilmu komputer yang menggabungkan statistik, machine learning, dan sistem basis data untuk menemukan pola yang berguna dalam kumpulan data besar.

#### **2. DDoS ICMP Flood**

DDoS ICMP Flood adalah jenis serangan DDoS yang memanfaatkan paket ICMP (Internet Control Message Protocol) untuk membanjiri target dengan permintaan echo (ping), menyebabkan kelebihan beban dan potensi kegagalan layanan.

#### **3. Dataset**

- Benign Traffic: berisi data lalu lintas jaringan yang normal.
- DDoS ICMP Flood: berisi data dari lalu lintas yang mengandung serangan ICMP flood.

## BAB III

### PEMBAHASAN

Analisis dilakukan menggunakan Python dan library pandas. Berikut langkah-langkah analisis:

Kode:

```
import pandas as pd
```

---

Kode:

```
dataset = pd.read_csv("Benign Traffic.csv")
```

---

Kode:

```
dataset
```

---

Kode:

```
dataset2 = pd.read_csv("DDoS ICMP Flood.csv")
```

---

Kode:

```
dataset
```

---

Kode:

```
dataset3= pd.read_csv("DDoS UDP Flood.csv")
```

---

Kode:

```
dataset3
```

---

Kode:

```
hasilgabung  
=pd.concat([dataset,dataset2,dataset3],ignore_index=True)
```

---

Kode:

```
hasilgabung.columns.values
```

---

Kode:

```
x = hasilgabung.iloc[:,7: 76]
```

---

Kode:

```
x
```

---

Kode:

```
y = hasilgabung.iloc[:,83: 84]
```

---

Kode:

```
y
```

---

Kode:

```
from sklearn.model_selection import train_test_split
```

---

Kode:

```
x_train,x_test,y_train,y_test = train_test_split(x,y, test_size = 0.2,  
random_state = 42)
```

---

Kode:

```
from sklearn import tree  
from sklearn.tree import DecisionTreeClassifier
```

---

Kode:

```
alya = DecisionTreeClassifier(criterion='entropy', splitter =  
'random')  
alya.fit(x_train,y_train)
```

---

Kode:

```
y_pred = alya.predict(x_test)
```

---

Kode:

```
y_pred
```

---

Kode:

```
from sklearn.metrics import accuracy_score  
accuracy = accuracy_score(y_test,y_pred)
```

---

Kode:

```
accuracy
```

---

Kode:

```
import matplotlib.pyplot as plt  
import numpy as np
```

---

Kode:

```
fig = plt.figure(figsize = (10, 7))  
tree.plot_tree(alya, feature_names = x.columns.values, class_names =  
np.array([ 'Benign Traffic','DDos ICMP Flood','DDoS UDP Flood']),  
filled = True)  
plt.show()
```

---

Kode:

```
import seaborn as lol  
from sklearn import metrics  
label = np.array([ 'Benign Traffic','DDos ICMP Flood','DDoS UDP  
Flood'])
```

---

Kode:

```
import matplotlib.pyplot as plt  
conf_matrix = metrics.confusion_matrix(y_test, y_pred)  
plt.figure(figsize=(10, 10))  
lol.heatmap(conf_matrix, annot=True, xticklabels=label,  
yticklabels=label)  
plt.xlabel('Prediksi')  
plt.ylabel('Fakta')  
plt.show()
```

---

Kode:

---

## BAB IV

### KESIMPULAN

Dari hasil analisis dua dataset lalu lintas jaringan, dapat disimpulkan bahwa proses data mining sangat berguna dalam membedakan antara lalu lintas normal dan lalu lintas yang mengandung serangan. Dengan menggunakan Python dan pustaka data seperti pandas, analisis data dapat dilakukan secara efisien untuk tujuan deteksi dan klasifikasi.

### LAMPIRAN

Berikut adalah kode lengkap yang digunakan dalam analisis:

```
import pandas as pd

dataset = pd.read_csv("Benign Traffic.csv")

dataset

dataset2 = pd.read_csv("DDoS ICMP Flood.csv")

dataset

dataset3= pd.read_csv("DDoS UDP Flood.csv")

dataset3

hasilgabung
=pd.concat([dataset,dataset2,dataset3],ignore_index=True)

hasilgabung.columns.values

x = hasilgabung.iloc[:,7: 76]

x

y = hasilgabung.iloc[:,83: 84]

y

from sklearn.model_selection import train_test_split

x_train,x_test,y_train,y_test = train_test_split(x,y, test_size = 0.2,
random_state = 42)
```

```
from sklearn import tree
from sklearn.tree import DecisionTreeClassifier

alya = DecisionTreeClassifier(criterion='entropy', splitter =
'random')
alya.fit(x_train,y_train)

y_pred = alya.predict(x_test)

y_pred

from sklearn.metrics import accuracy_score
accuracy = accuracy_score(y_test,y_pred)

accuracy

import matplotlib.pyplot as plt
import numpy as np

fig = plt.figure(figsize = (10, 7))
tree.plot_tree(alya, feature_names = x.columns.values, class_names =
np.array([ 'Benign Traffic','DDos ICMP Flood','DDoS UDP Flood']),
filled = True)
plt.show()

import seaborn as lol
from sklearn import metrics
label = np.array([ 'Benign Traffic','DDos ICMP Flood','DDoS UDP
Flood'])

import matplotlib.pyplot as plt
conf_matrix = metrics.confusion_matrix(y_test, y_pred)
plt.figure(figsize=(10, 10))
lol.heatmap(conf_matrix, annot=True, xticklabels=label,
yticklabels=label)
plt.xlabel('Prediksi')
plt.ylabel('Fakta')
plt.show()
```

---