



## Lembar Kerja Mahasiswa

Mata Kuliah : Data Mining  
Bahasan : Ansamble Classification  
Halaman : 1/3

NIM	222410101087
Nama	Rafi Jauhari
Kelas	A
Program Studi	Sistem Informasi
Asisten	1. Renata Sayidatul Arikha 212410101057 2. Aprodhita Nanda Eka Wijaya 212410101071

### LANGKAH KERJA

1. Buat proses klasifikasi menggunakan metode Boosting khusus untuk data kategori pada dataset adult.csv. Tuliskan penjelasan analisis dan akurasi.

```
import pandas as pd
import numpy as np
from sklearn.ensemble import AdaBoostClassifier
from sklearn.tree import DecisionTreeClassifier
from sklearn.preprocessing import LabelEncoder, StandardScaler
from sklearn.model_selection import train_test_split
from sklearn.metrics import accuracy_score

#Membaca dataset adult
data = pd.read_csv("adult.csv")
data.head()

#Menghilangkan baris yang ada nilai kosong
data = data.dropna()

#Memberi label encoder kepada variabel yang bersifat kategori
label_encoders = {}
for column in data.select_dtypes(include=['object']).columns:
    le = LabelEncoder()
    data[column] = le.fit_transform(data[column])
    label_encoders[column] = le
```



## Lembar Kerja Mahasiswa

Mata Kuliah : Data Mining  
Bahasan : Ensemble Classification  
Halaman : 2/3

```
#Membuat target
X = data.drop('income', axis=1)
y = data['income']

#Standarisasi
scaler = StandardScaler()
X = scaler.fit_transform(X)

#Membuat data training sets
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2,
random_state=42)

#Membuat klasifikasi
base_estimator = DecisionTreeClassifier(max_depth=1)
adaboost = AdaBoostClassifier(base_estimator=base_estimator, n_estimators=50,
learning_rate=1.0, random_state=42)

#Melatih
adaboost.fit(X_train, y_train)

#Membuat Prediksi
y_pred = adaboost.predict(X_test)

#Total Akurasi
accuracy = accuracy_score(y_test, y_pred)
print(f'Accuracy: {accuracy:.2f}')
```

Hasil :

**Accuracy: 0.86**

### HASIL DAN ANALISIS DATA

Dari proses klasifikasi menggunakan metode Boosting khusus untuk data kategori pada dataset adult.csv yang saya lakukan dengan target dari klasifikasi adalah variabel income di dataset adult.csv saya mendapatkan akurasi yang lumayan tinggi yaitu sebesar 0,86.



## Lembar Kerja Mahasiswa

Mata Kuliah : Data Mining  
Bahasan : Ansamble Classification  
Halaman : 3/3

### KESIMPULAN

Dari hasil ini, dapat disimpulkan bahwa model memiliki performa yang baik dalam mengklasifikasikan data dengan akurasi tinggi, namun performa untuk kelas 1 sedikit lebih rendah yang bisa jadi karena ketidakseimbangan kelas atau kompleksitas data.

Link Google Colab	<a href="https://colab.research.google.com/drive/1rhtrJXzda9i5MMIAykCYB6REvjxZRLcJ#scrollTo=kWPOMePQNcV6">https://colab.research.google.com/drive/1rhtrJXzda9i5MMIAykCYB6REvjxZRLcJ#scrollTo=kWPOMePQNcV6</a>
-------------------	---

Link Youtube (Unlisted)	
-------------------------	--

Jember, .....2024

Mengetahui,  
Dosen Datamining

Asisten,

Fajrin Nurman Arifin, S.T., M.Eng  
NIP. 198511282015041002

(Nama Jelas)  
NIM.