

# **Proyek Mesin Learning Prediksi Kelulusan Mahasiswa Berdasarkan Data Akademik**



Disusun Oleh :

Nama : Nisrina Zulfa R  
NIM : A11.2023.15076  
Kelas : A11.4402  
Dosen Pengampu : Pak Abu Salam M.Kom

Link Git-Hub : <https://github.com/nisrina-26/ML-A11.2023.15076-UAS>

UNIVERSITAS DIAN NUSWANTORO  
SEMARANG  
2025

## Ringkasan dan Permasalahan

Permasalahan yang diangkat dalam proyek ini adalah bagaimana memprediksi apakah seorang mahasiswa akan lulus tepat waktu atau tidak berdasarkan data akademik dan demografis. Permasalahan ini penting untuk mendukung proses monitoring dan evaluasi akademik.

Banyak perguruan tinggi menghadapi tantangan dalam memantau dan memprediksi kelulusan mahasiswa secara tepat waktu. Keterlambatan dalam lulus dapat berdampak pada efisiensi institusi, beban keuangan mahasiswa, serta reputasi akademik kampus. Namun, tidak semua institusi memiliki alat analitik yang mampu mengidentifikasi mahasiswa yang berisiko terlambat lulus atau tidak lulus.

## Tujuan

- Memprediksi kelulusan mahasiswa dengan akurasi tinggi.
- Menyediakan alat bantu analisis bagi pihak akademik.
- Membangun model prediksi menggunakan data akademik mahasiswa (seperti IPK, SKS, nilai mata kuliah inti, jumlah pengambilan semester, dan keaktifan studi)

## Manfaat:

- Memberikan *early warning system* kepada pihak akademik.
- Membantu dosen wali atau pembimbing dalam pemantauan kemajuan studi mahasiswa.
- Meningkatkan tingkat kelulusan tepat waktu.
- Mengoptimalkan strategi intervensi dan bimbingan.

## Penyebab Umum Ketidakihtulusan Tepat Waktu:

- IPK rendah atau tidak stabil.
- Jumlah SKS lulus yang tidak mencukupi tiap semester.
- Sering mengulang mata kuliah inti.
- Ketidakhadiran tinggi atau masa cuti studi.
- Masalah non-akademik (yang dalam konteks ini diasumsikan tidak tersedia dalam dataset).

## Model / Alur Penyelesaian

graph TD

A[Data Collection] --> B[Data Cleaning]

B --> C[Exploratory Data Analysis]

C --> D[Feature Engineering]

D --> E[Modeling (Training & Testing)]

E --> F[Evaluation]

F --> G[Kesimpulan]

## Dataset

Dataset berisi data akademik mahasiswa yang mencakup:

- IPK
- Jumlah SKS yang diambil
- Kehadiran
- Jenis Kelamin
- Asal Sekolah
- Status Lulus (0 = Tidak, 1 = Lulus)

```
import pandas as pd
import seaborn as sns
import matplotlib.pyplot as plt

# Load dataset
df = pd.read_csv("mahasiswa.csv")
df.head()

# Informasi Umum Dataset
df.info()
```

```
# Cek nilai yang hilang
df.isnull().sum()
```

```
# Statistik Deskriptif
df.describe()
```

## EDA (Exploratory Data Analysis)

```
# Distribusi Status Kelulusan
sns.countplot(x="Status_Lulus", data=df)
plt.title("Distribusi Kelulusan Mahasiswa")
plt.show()

# Korelasi antar Fitur
sns.heatmap(df.corr(), annot=True)
plt.title("Heatmap Korelasi Fitur")
plt.show()
```

## Feature Engineering

```
# Feature Engineering
df['Gender'] = df['Gender'].map({'Pria': 1, 'Wanita': 0})
df = pd.get_dummies(df, columns=['Asal_Sekolah'], drop_first=True)
df.head()
```

## Proses Learning / Modeling

```
from sklearn.model_selection import train_test_split
from sklearn.ensemble import RandomForestClassifier
from sklearn.metrics import classification_report, confusion_matrix, accuracy_score

# Definisi fitur dan target
X = df.drop('Status_Lulus', axis=1)
y = df['Status_Lulus']

# Split dataset
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=42)

# Training model
model = RandomForestClassifier(random_state=42)
model.fit(X_train, y_train)

# Prediksi
y_pred = model.predict(X_test)
```

## Performa Model

```
print("Confusion Matrix:\n", confusion_matrix(y_test, y_pred))
print("\nClassification Report:\n", classification_report(y_test, y_pred))
print("\nAccuracy Score:", accuracy_score(y_test, y_pred))
```

## Diskusi Hasil dan Kesimpulan

### Diskusi Hasil

Model Random Forest menunjukkan akurasi tinggi dalam memprediksi kelulusan mahasiswa. Fitur paling berpengaruh adalah IPK, jumlah SKS, dan kehadiran.

### Kesimpulan

- Model berhasil melakukan prediksi kelulusan dengan baik.
- Dapat dijadikan alat bantu oleh pihak akademik untuk memantau kemajuan mahasiswa.