

**MACHINE LEARNING**

**LAPORAN MODEL KLASIFIKASI PREDIKSI KELULUSAN**



**MUHAMMAD ADITYA SAPUTRA      231011400243**

**PROGRAM STUDI TEKNIK INFORMATIKA S-1**

**FAKULTAS ILMU KOMPUTER**

**UNIVERSITAS PAMULANG**

**TAHUN 2025**

## 1. DESKRIPSI DATASET

Dataset yang digunakan dalam proyek ini adalah Iris Dataset, yang disimulasikan menjadi data prediksi kelulusan berdasarkan beberapa atribut seperti:

- Sepal Length
- Sepal Width
- Petal Length
- Petal Width

Target variabelnya adalah kelas bunga (Setosa, Versicolor, Virginica) yang diibaratkan sebagai kategori kelulusan.

Jumlah data: 150 baris

Jumlah fitur: 4

Tipe target: Kategorikal (3 kelas)

	sepal length (cm)	sepal width (cm)	petal length (cm)	petal width (cm)	kelulusan
0	5.1	3.5	1.4	0.2	Tidak Lulus
1	4.9	3.0	1.4	0.2	Tidak Lulus
2	4.7	3.2	1.3	0.2	Tidak Lulus
3	4.6	3.1	1.5	0.2	Tidak Lulus
4	5.0	3.6	1.4	0.2	Tidak Lulus

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 150 entries, 0 to 149
Data columns (total 5 columns):
#   Column                Non-Null Count  Dtype
---  -
0   sepal length (cm)      150 non-null   float64
1   sepal width (cm)       150 non-null   float64
2   petal length (cm)      150 non-null   float64
3   petal width (cm)       150 non-null   float64
4   kelulusan              150 non-null   object
dtypes: float64(4), object(1)
memory usage: 6.0+ KB

kelulusan
Lulus      100
Tidak Lulus  50
Name: count, dtype: int64
```

## 2. MODEL YANG DIGUNAKAN

Dalam proyek ini digunakan dua model klasifikasi:

### 1. Logistic Regression

Model berbasis regresi logistik yang cocok untuk klasifikasi multikelas dengan probabilitas.

Accuracy: 1.0

Confusion Matrix:

```
[[20  0]
 [ 0 10]]
```

Classification Report:

	precision	recall	f1-score	support
--	-----------	--------	----------	---------

Lulus	1.00	1.00	1.00	20
Tidak Lulus	1.00	1.00	1.00	10
accuracy			1.00	30
macro avg	1.00	1.00	1.00	30
weighted avg	1.00	1.00	1.00	30

## 2. Decision Tree Classifier

Model berbasis pohon keputusan yang membagi data berdasarkan nilai atribut untuk membuat keputusan akhir.

Accuracy: 1.0

Confusion Matrix:

```
[[20  0]
 [ 0 10]]
```

Classification Report:

	precision	recall	f1-score	support
...				
accuracy			1.00	30
macro avg	1.00	1.00	1.00	30
weighted avg	1.00	1.00	1.00	30

## 3. KESIMPULAN

Berdasarkan hasil pengujian menggunakan dua algoritma machine learning, yaitu Logistic Regression dan Decision Tree, dapat disimpulkan hal-hal berikut:

1. Kedua model berhasil melakukan prediksi kelulusan mahasiswa dengan akurasi 100%. Hal ini menunjukkan bahwa data yang digunakan memiliki pola yang sangat mudah dipelajari oleh model.
2. Logistic Regression bekerja dengan baik pada data yang memiliki hubungan linear antara fitur dan target. Model ini cocok digunakan untuk data yang relatif sederhana dan terdistribusi secara seimbang.
3. Decision Tree menunjukkan hasil yang sama baiknya dan memiliki kelebihan karena dapat menangani data non-linear serta tidak memerlukan proses standarisasi. Model ini lebih fleksibel terhadap variasi data.
4. Hasil sempurna pada dataset ini tidak sepenuhnya menggambarkan performa di dunia nyata. Untuk aplikasi sebenarnya (misalnya prediksi kelulusan mahasiswa berdasarkan nilai, absensi, IPK, dan lama studi), diperlukan dataset yang lebih besar dan kompleks agar model dapat diuji secara realistis.
5. Dengan hasil ini, dapat disimpulkan bahwa algoritma Logistic Regression dan Decision Tree sama-sama efektif untuk klasifikasi sederhana, dan keduanya bisa dijadikan dasar untuk penelitian lanjutan menggunakan dataset akademik yang lebih lengkap.









