

MACHINE LEARNING
DOSEN : AGUNG PERDANANTO S.Kom, M.Kom.



FAKULTAS ILMU KOMPUTER
PROGRAM STUDI TEKNIK INFORMATIKA
UNIVERSITAS PAMULANG
2025

NAMA : NAQA ARVA DERINA

KELAS : 05TPLE005

NIM : 231011400221

UTS

LAPORAN

1. Pendahuluan

Penelitian ini bertujuan untuk membangun model klasifikasi menggunakan dataset Titanic. Tujuannya adalah untuk memprediksi apakah seorang penumpang selamat atau tidak berdasarkan atribut seperti umur, jenis kelamin, kelas tiket, dan tarif.

Model klasifikasi ini membantu memahami faktor-faktor yang mempengaruhi keselamatan penumpang.

2. Dataset

Dataset yang digunakan berasal dari GitHub:

<https://github.com/datasciencedojo/datasets/blob/master/titanic.csv> Dataset

berisi 891 data penumpang dengan beberapa fitur penting:

- **Pclass:** Kelas tiket (1, 2, 3)
- **Sex:** Jenis kelamin
- **Age:** Usia penumpang
- **Fare:** Biaya tiket
- **Embarked:** Pelabuhan keberangkatan
- **Survived:** 1 = Selamat, 0 = Tidak selamat (target)

3. Metodologi

Langkah-langkah yang dilakukan:

1. **Data Cleaning:** Menghapus kolom tidak relevan (Name, Ticket, Cabin) dan mengisi nilai kosong.
2. **Encoding:** Mengubah data kategorikal menjadi numerik (Sex, Embarked).
3. **Normalisasi:** Standarisasi fitur numerik menggunakan StandardScaler.
4. **Pemodelan:** Menggunakan dua algoritma:
 - Logistic Regression
 - Decision Tree
5. **Evaluasi:** Mengukur performa model dengan akurasi, presisi, recall, dan F1-score.

4. Hasil Evaluasi

Contoh hasil evaluasi (bisa diganti sesuai hasil dari notebook kamu):

Model	Accuracy	Precision	Recall	F1-Score
Logistic Regression	0.80	0.78	0.74	0.76
Decision Tree	0.83	0.80	0.79	0.79

5. Kesimpulan

Berdasarkan hasil evaluasi, model **Decision Tree** memiliki performa yang sedikit lebih baik dibandingkan Logistic Regression.

Hal ini menunjukkan bahwa Decision Tree mampu menangkap pola hubungan antar fitur dengan lebih kompleks.

Namun, Logistic Regression tetap unggul dalam hal interpretabilitas dan efisiensi.

Model ini dapat dikembangkan lebih lanjut dengan teknik tuning hyperparameter dan crossvalidation untuk meningkatkan akurasi.