

Report for Titanic Dataset

A. Feature Classification

Numerical: Age, Fare

Categorical: Embarked

Ordinal: Pclass (Represents socio-economic status: 1 > 2 > 3)

Binary: Survived (Target), Sex

B. Data Quality Observations

Missing Values: The Age column has missing entries that need imputation, and the Cabin column is mostly empty.

Data Imbalance: The number of "Deceased" (0) is significantly higher than "Survived" (1), which is a key consideration for Machine Learning.

C. ML Suitability

The dataset is highly suitable for binary classification. The Survived column is the target variable.

Features like Sex, Pclass, and Fare show strong correlations with survival rates, making them excellent input features for a model.