Titanic Survival Prediction — Final Report

Prepared By: [Your Name]

Date: [Submission Date]

# Table of Contents

1. 1. Introduction
2. 2. Data Analysis (EDA)
3. 3. Data Preprocessing
4. 4. Model Building
5. 5. Accuracy Evaluation
6. 6. Conclusion
7. 7. Attachments

# 1. Introduction

The Titanic dataset is a classic machine learning problem...

# 2. Data Analysis (EDA)

We analyzed the training data (train.csv) and observed the following key points: - Gender: Females had a much higher survival rate compared to males. - Passenger Class (Pclass): First-class passengers had higher survival rates. - Age: Younger passengers had better survival chances. - Family Size: Minor effects observed. - Missing Data: In Age, Cabin, Embarked columns.

# 3. Data Preprocessing

Selected Features: Pclass, Sex, Age, SibSp, Parch, Fare. Handled Missing Values: Filled missing Age values. Encoded Categorical Data: Mapped 'Sex' to numbers.

# 4. Model Building

Used Decision Tree Classifier. Training was done using selected features. Generated submission file with predictions.

# 5. Accuracy Evaluation

Training Accuracy: ~98% Cross-validation Mean Accuracy: ~79% (Model shows slight overfitting).

# 6. Conclusion

Performed EDA, preprocessing, modeling, and prediction. Achieved decent results with Decision Tree. Future improvements: feature engineering, better models.

# 7. Attachments

submission.csv (ready for upload or evaluation)