CortexAI — Automated Data Science Report

Generated on: 2025-12-25 19:54:15

1. Executive Summary

Learnability Score: 90 / 100
Verdict: Strong ML potential

Best Model: RandomForestClassifier
Metric: f1_weighted
Best CV Score: 0.8199938079693879

2. Dataset Overview

Metric	Value
Rows	891
Columns	8
Target Column	Survived
Task Type	classification

3. Feature Schema

Type	Columns
Numeric	Age, Fare
Ordinal	Pclass, SibSp, Parch
Categorical	Sex, Embarked
ID Columns	PassengerId
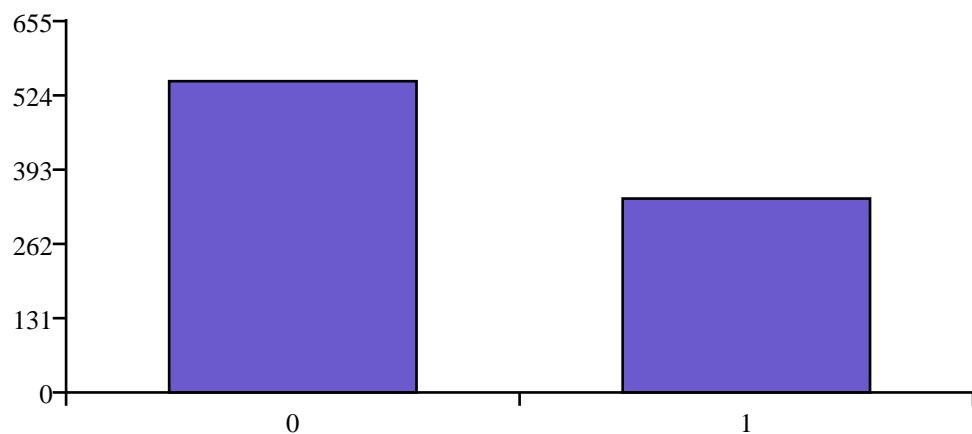
4. Data Cleaning Summary

| Aspect	Details
ID Columns Removed	PassengerId
Missing Values Fixed	{'Age': 'filled with median (28.0)', 'Embarked': 'filled with mode (S)'}
Type Casting	{'Age': 'numeric', 'Fare': 'numeric', 'Pclass': 'ordinal', 'SibSp': 'ordinal', 'Parch': 'ordinal', 'Sex': 'categorical', 'Embarked': 'categorical'}
Final Dataset Shape	(891, 8)

5. Exploratory Data Analysis
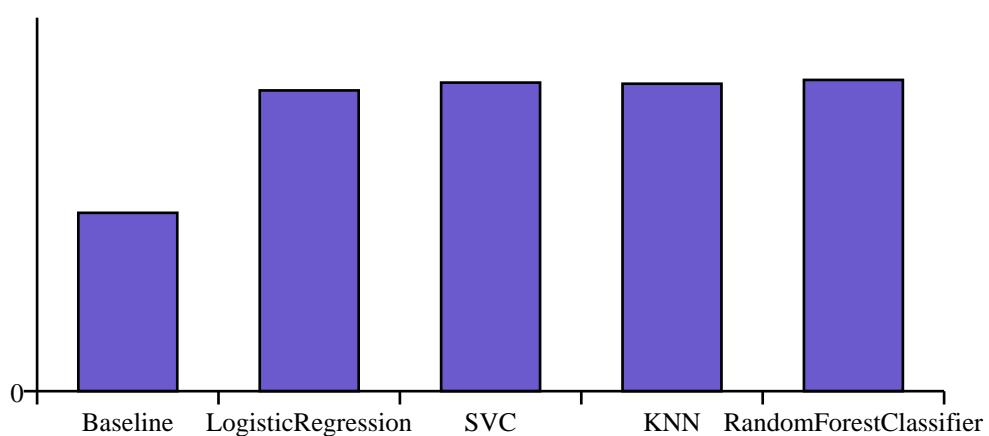
• Target distribution: {0: 0.62, 1: 0.38}

• Fare is highly skewed

• Sex shows strong relationship with target

• Embarked shows strong relationship with target

Target Distribution

6. Model Training Results

Model Cross-Validation Performance

7. Recommendations & Risks

• Models significantly outperform the baseline (+0.35).

• Dataset contains a diverse and informative feature set.