CortexAI — Automated Data Science Report

Generated on: 2025-12-26 00:53

This report was automatically generated by CortexAI to assess dataset quality, perform exploratory data analysis, train multiple machine learning models, and provide actionable insights.

Learnability Score: 80 / 100

Verdict: Strong ML potential

1. Dataset Overview

Rows	150
Columns	5
Target Column	Species
Task Type	classification

2. Feature Schema

Numeric Features	SepalLengthCm, SepalWidthCm, PetalLengthCm, PetalWidthCm
Ordinal Features	
Categorical Features	Species
Datetime Features	
ID Columns	Id

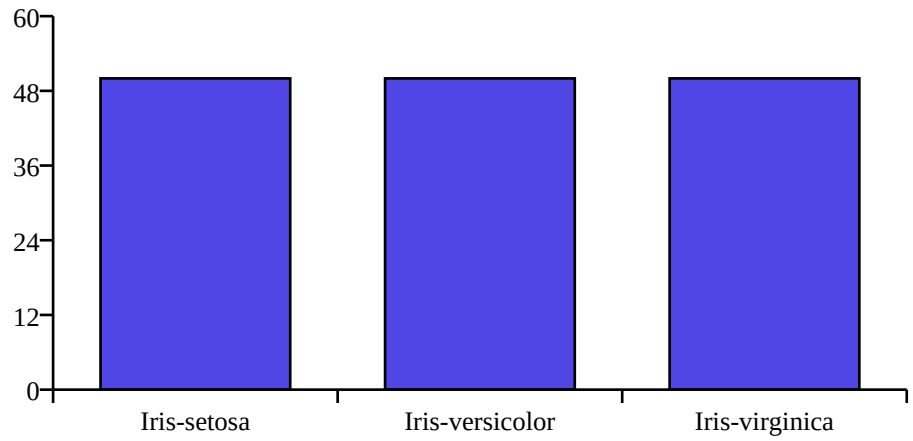
3. Data Cleaning Summary

Dropped ID Columns	Id
Dropped High Cardinality	
Missing Values Fixed	
Type Casting	SepalLengthCm: numeric, SepalWidthCm: numeric, PetalLengthCm: numeric, PetalWidthCm: numeric
Final Dataset Shape	150 × 5

4. Exploratory Data Analysis

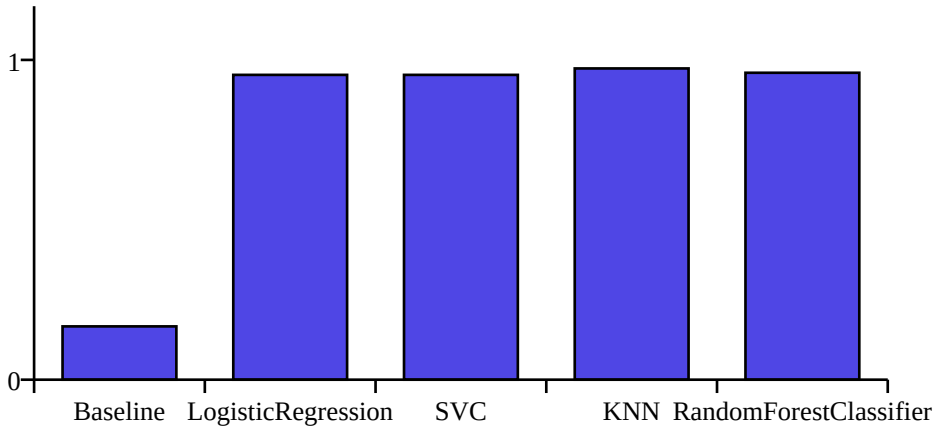
• Target distribution: {'Iris-setosa': 0.33, 'Iris-versicolor': 0.33, 'Iris-virginica': 0.33}

Target Distribution



5. Model Performance

Model Comparison



Best Model: KNN

Best Score: 0.9732999164578111 (f1_macro)

6. Strengths

✓ Models significantly outperform the baseline (+0.81).

✓ Dataset is clean with no missing values.

✓ Target classes are well-balanced.

✓ Features are numerically well-behaved and separable.

7. Risks & Limitations

No significant data quality risks or modeling limitations were identified for this dataset.

8. Recommendations

No critical corrective actions are required. The dataset is suitable for machine learning in its current form.