

The Higgs Boson Challenge

Project By:

1. Ashish Kumar
2. Chandra Prakash
3. Karthik Rathod
4. Rashmi Deshpande
5. Yamuna Katta

Problem Statement:

The ATLAS experiment has recently observed a signal of the Higgs boson decaying into two tau particles, but this decay is a small signal buried in background noise.

The goal of the project is to improve the procedure that produces the selection region.



What is Higgs
Boson???



①

Preprocessing
& Cleaning
Data

③

Base Model Execution

④

Ensemble Models

⑤

Exploratory Data Analysis

Experimentation &
Interpretation

Upscaling The Data

Transforming The
Data

Rectifying Colinearity

Data Analysis

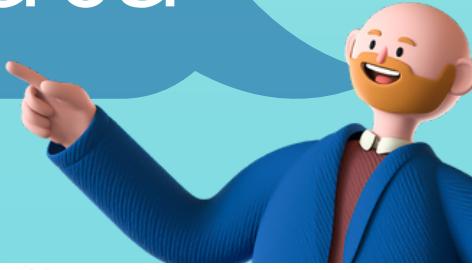
Data Profiling

Data Extraction

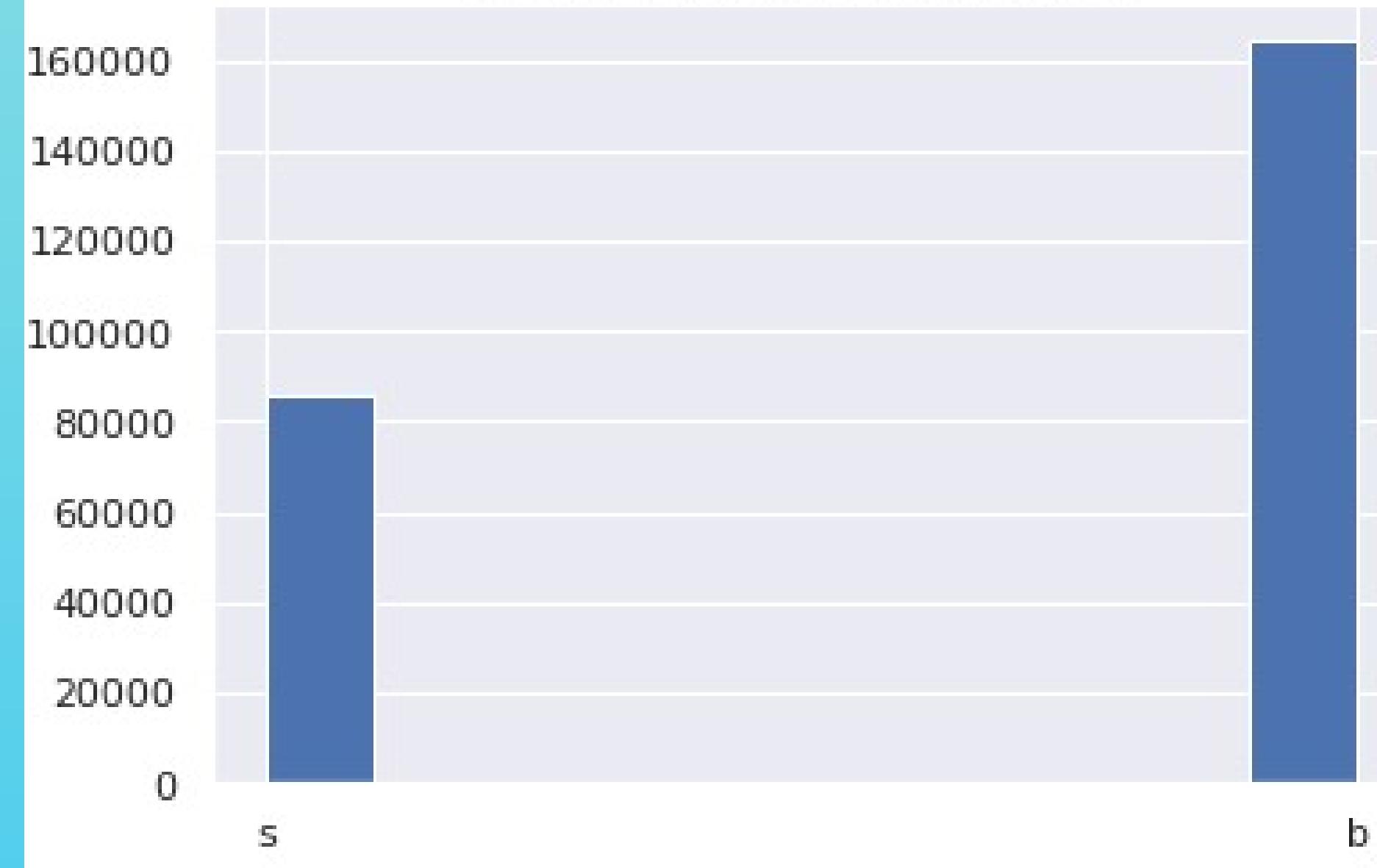


How Did We
Clean the Data?

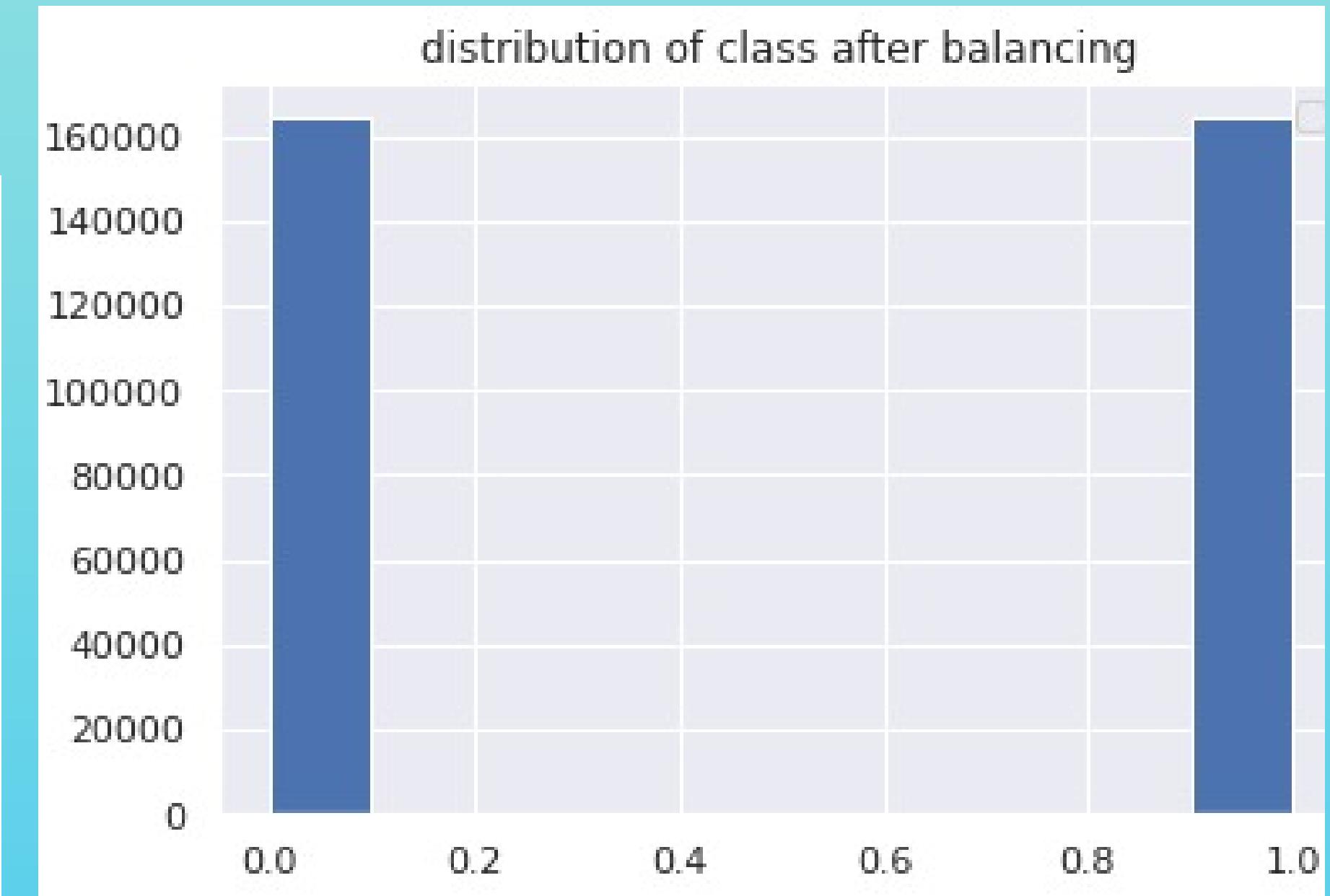
Preprocessed Data



distribution of class in data set



distribution of class after balancing



Cleaned Data



Decision Tree
Classifier

Random Forest
Classifier

Gradient Boosting
Classifier

M
O
D
E
L
S

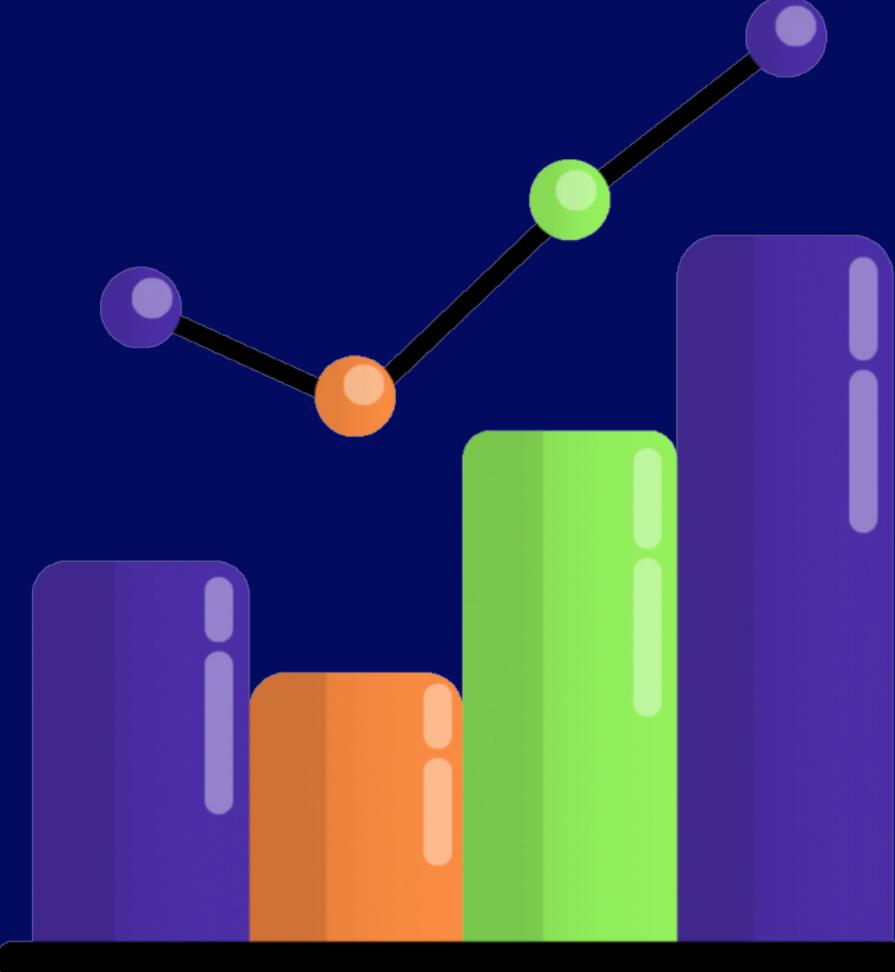
Logistic
Regression

Bagging
Classifier

MODEL PERFORMANCE



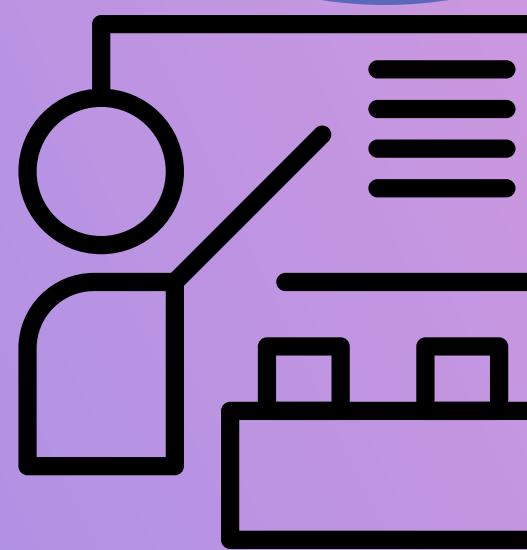
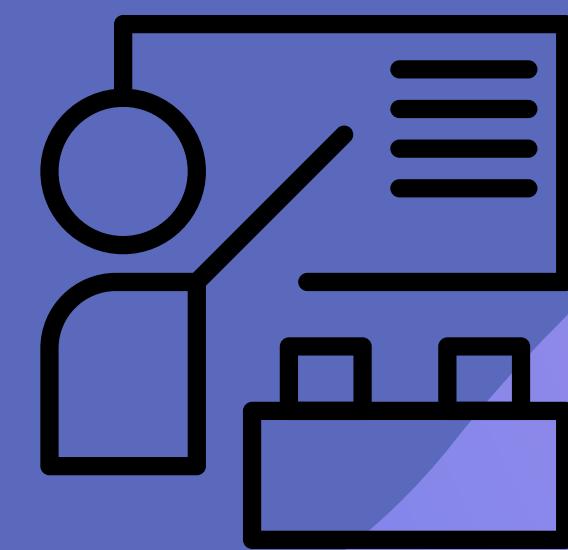
RESULTS



Random Forest Classification

ACCURACY	0 . 68
F1 Score	0 . 08
ROC Score	0 . 50

Conclusion



The Random Forest Model performance is better than others.

Model	F1 Score	ROC Score	Accuracy
Logistic Regression	0.41	0.50	0.43
Bagging	0.17	0.50	0.65
Decision Tree	0.17	0.49	0.65
Random Forest	0.08	0.50	0.68

Future Scope

Better data preprocessing,
Try different methods for
improving the data quality

Improving scores by
performing hyper
parameter tuning

Trying addaboost for
better scores