



IronHack Personal Project

Home Loan Model Development

GitHub Repository:
https://github.com/MC993/home_loan_approval

Mirko Caroleo
28.03.2025





Agenda

- 1. Scope of the project**
- 2. Process**
- 3. Results**
- 4. Streamlit Presentation**
- 5. Q&A**



1. Scope of the Project

- The dataset provided consists of historical data from a Loan Company.
- The task is to create a model which is able to predict whether an applicant should be granted a loan or not based on important features.
- I achieved positive results via:
 - EDA & Data Cleaning,
 - Feature Engineering,
 - Hyperparameters Tuning.



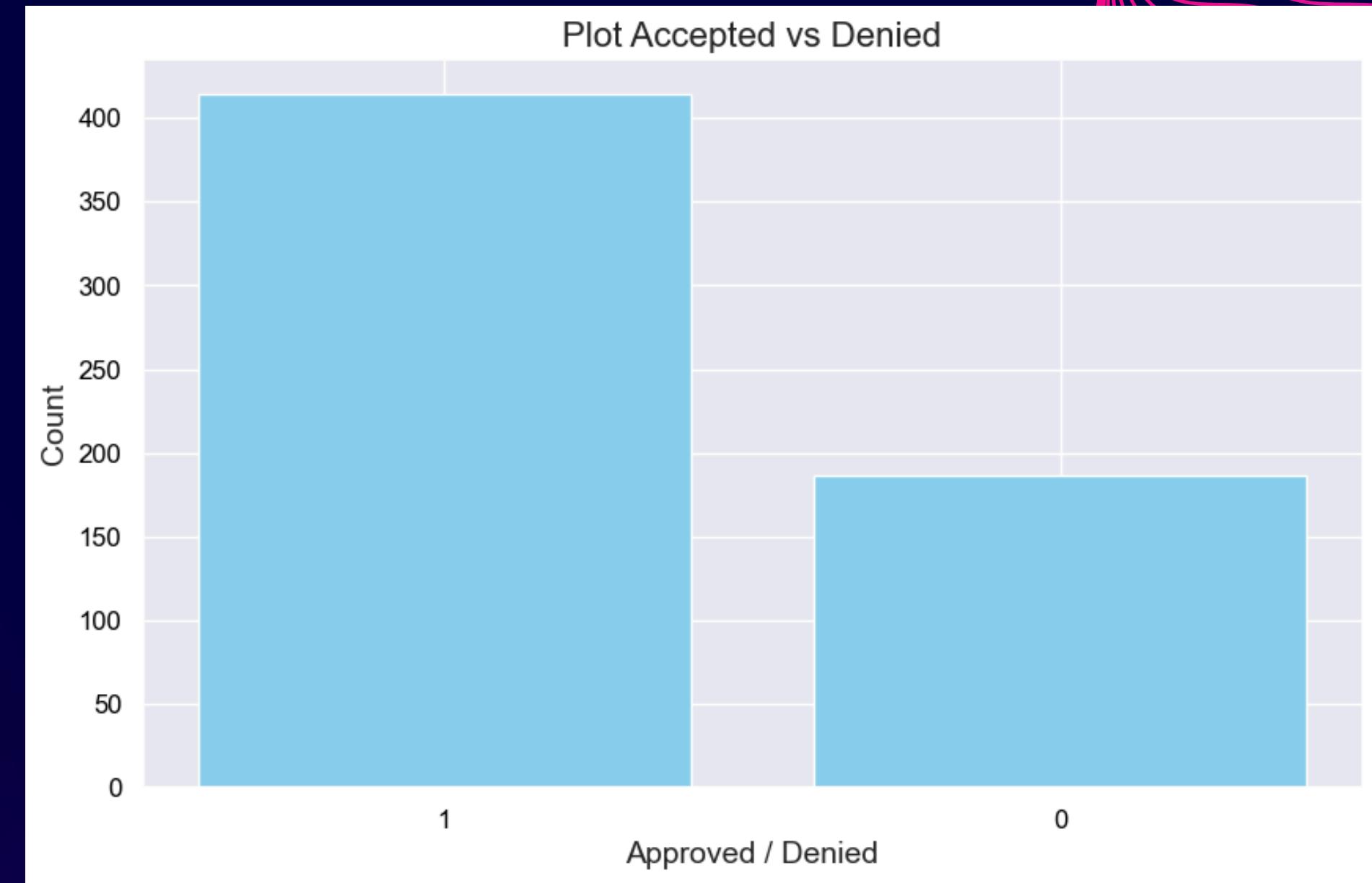
2. Process

- **EDA & Data Cleaning**
- **Creation of a baseline model**
- **Feature Engineering**
- **Hyperparameters analysis**



EDA & Data Analysis

- Filled missing values
- Checked the distribution
- Checked potential mismatch between Approved and Denied data
- Resampled
- code snipped to resample:



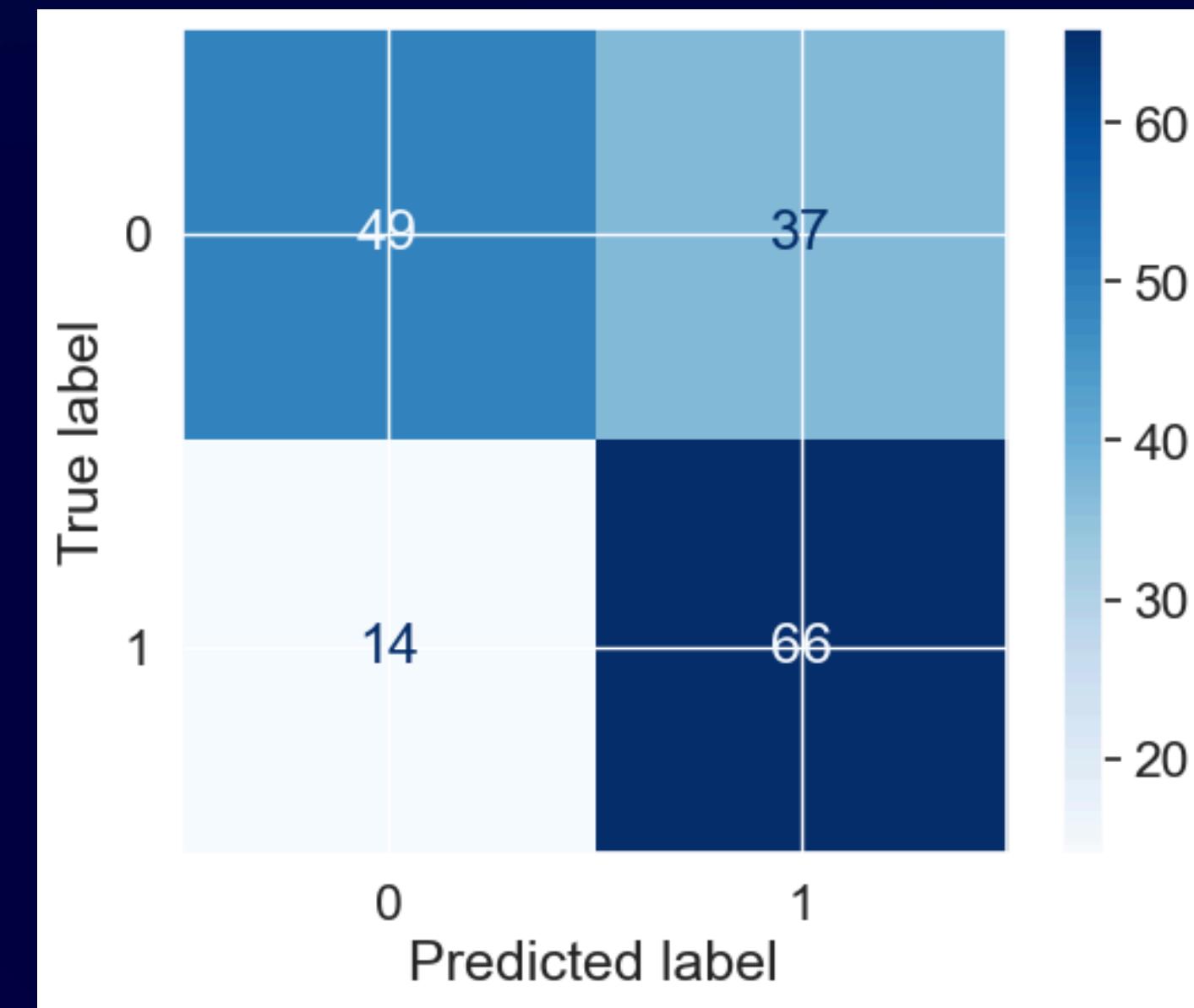
```
from sklearn.utils import resample  
no_approval = train[train['loan_Y']==0]  
approval = train[train['loan_Y']==1]  
oversampled_no_approval = resample(no_approval, replace=True, n_samples=len(approval),  
random_state=42)
```



IronHack Personal Project

Baseline Model (LogisticRegression)

- Average Precision on Denied: 78%
- Average Precision on Approved: 64%

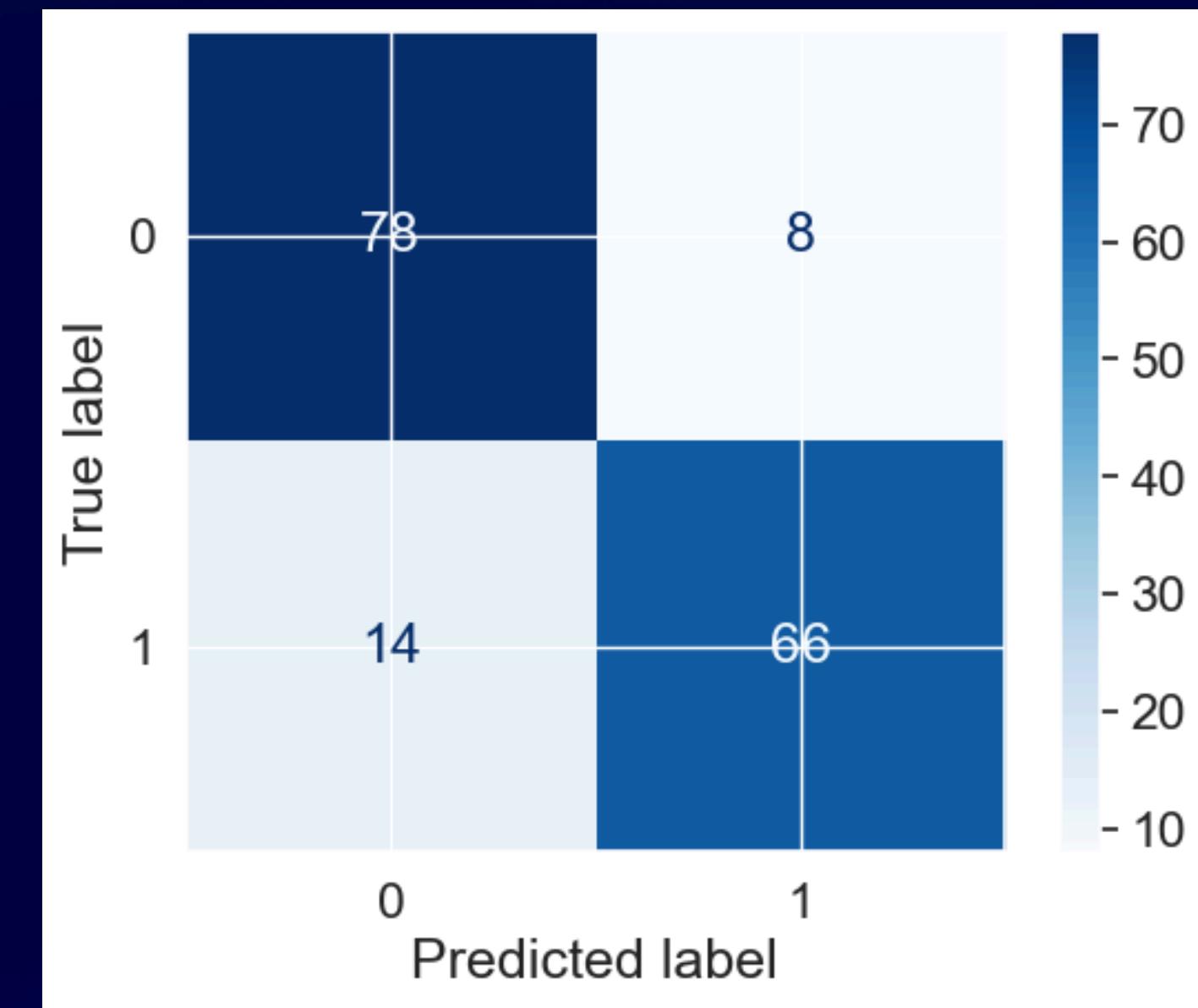


	precision	recall	f1-score	support
0	0.78	0.57	0.66	86
1	0.64	0.82	0.72	80
accuracy			0.69	166
macro avg	0.71	0.70	0.69	166
weighted avg	0.71	0.69	0.69	166



Hyperparameter Tuning (XGBClassifier)

- Average Precision on Denied: 85%
- Average Precision on Approved: 89%
- Best Hyperparameter selected: max_depth 20

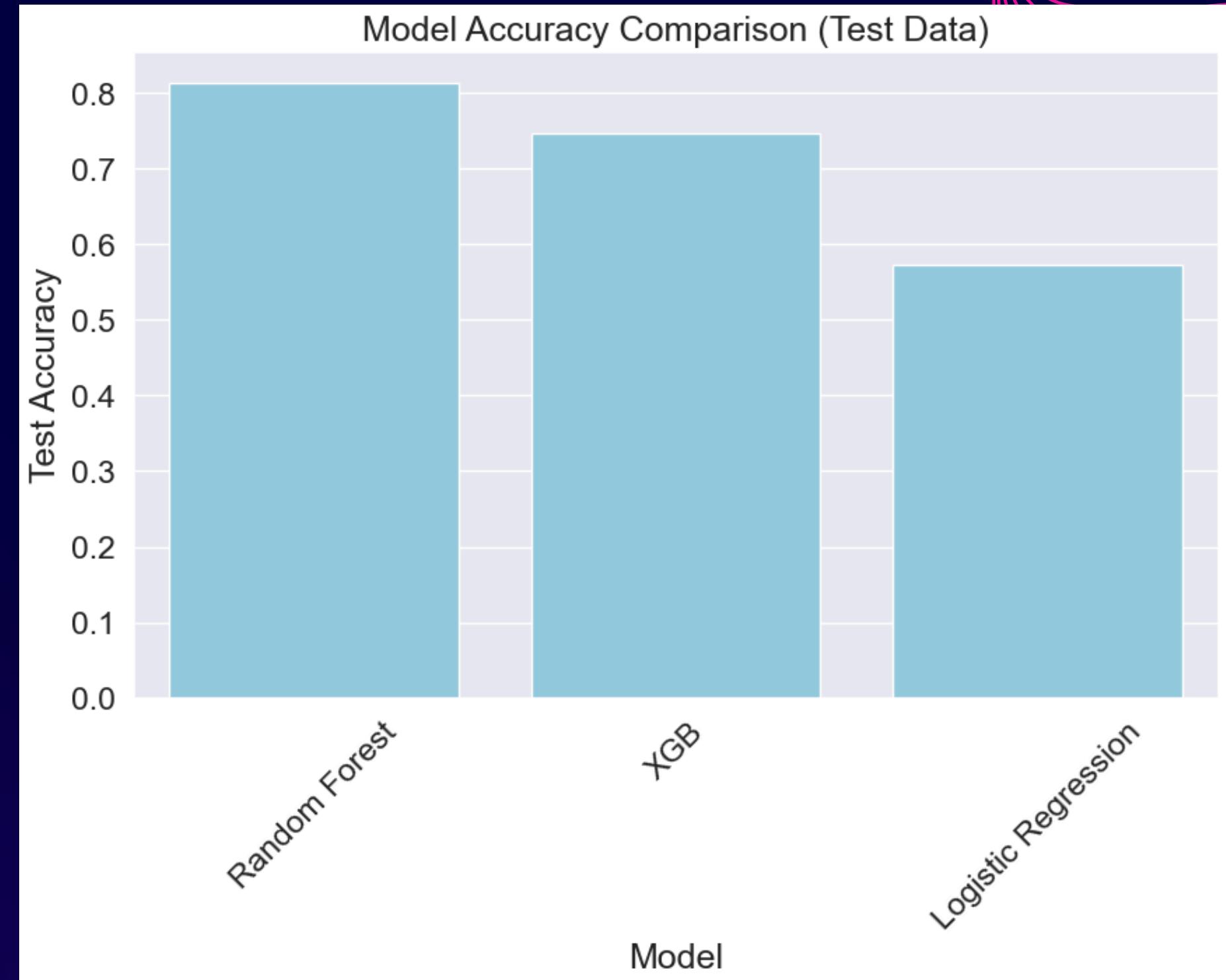


	precision	recall	f1-score	support
0	0.85	0.91	0.88	86
1	0.89	0.82	0.86	80
accuracy			0.87	166
macro avg	0.87	0.87	0.87	166
weighted avg	0.87	0.87	0.87	166



3. Results

- RandomForest
Average Test
Precision: 81%
- XGB Average Test
Precision: 74%
- LogisticRegression
Average Test
Precision: 57%





IronHack Personal Project

Q&A