Indian Bank Loan Approval Project .

PRESENTATION

Table Of Contents.

I Overview

In this project, machine learning methods are used to analyse data from earlier clients to identify trends that might be used to mark a candidate as a risky candidate or .

1.To achieve a target accuracy score of 80% for the predictive model, ensuring reliable predictions that aid in decisionmaking .

2..To create a machine learning model that can accurately predict the likelihood of loan approval 0 for applicants.

3.To use insights from the Model to automate the loan approval 03 system based on best features .

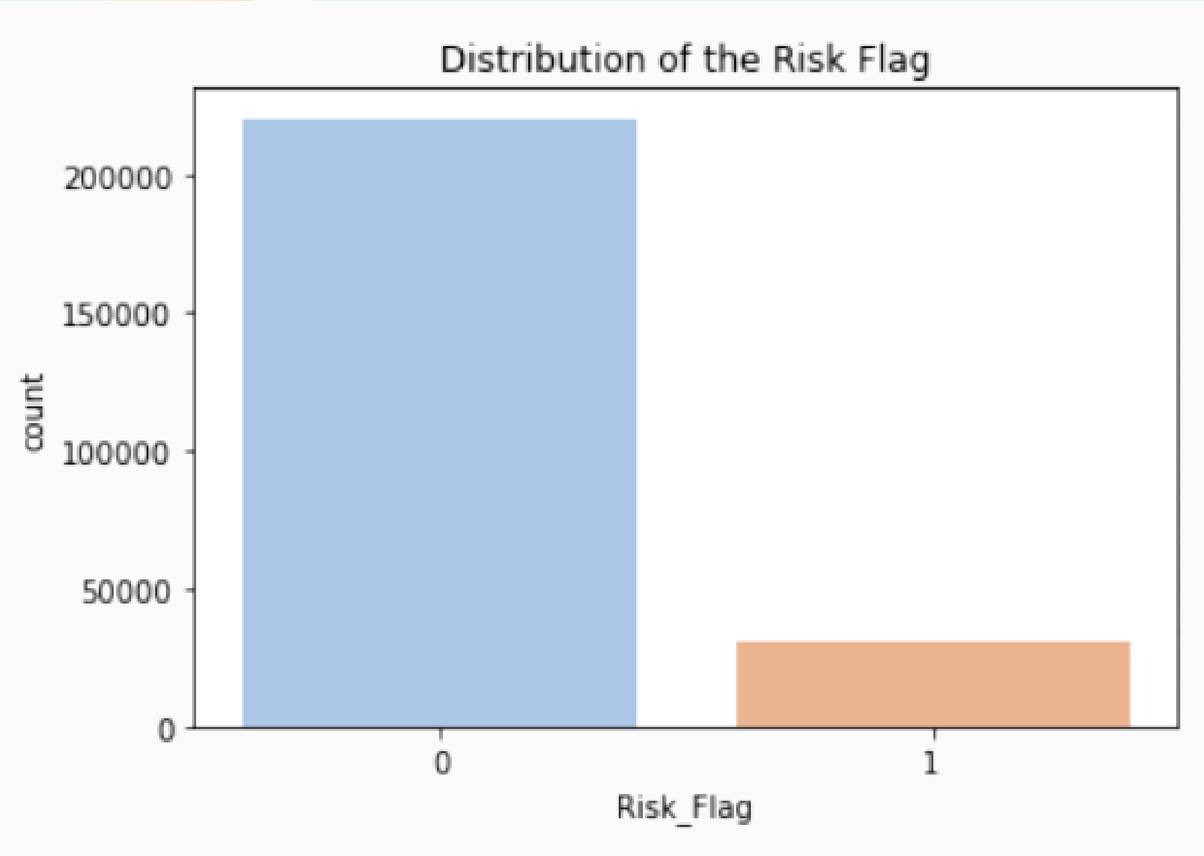
problem statement

The state bank of india wishes to enhance its loan approval process by developing a predictive model that accurately determines the likelihood of loan aprroval for applicants

Data Understanding.

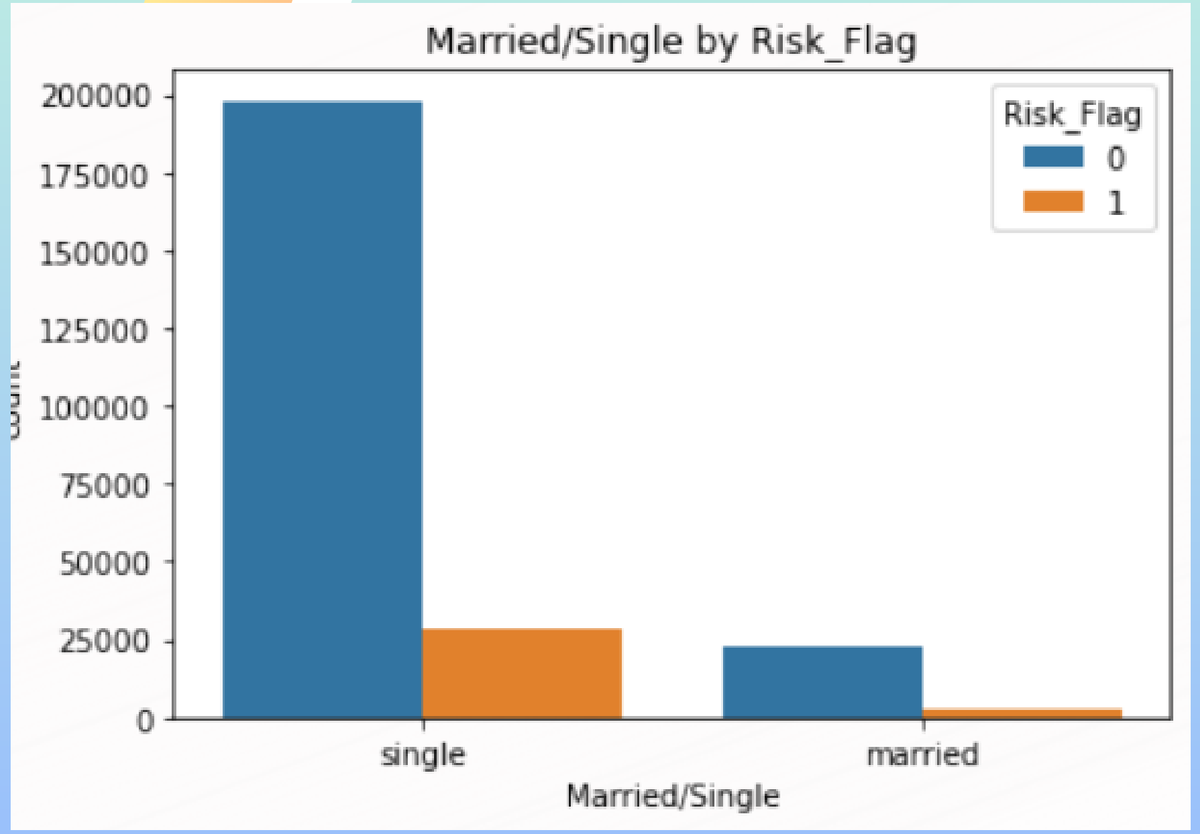
The loan approval dataset was obtained from Kaggle and contains extensive details about a loan applicant's financial status and personal characteristics, including age and marital status.

Analysis of the Risk Flag information



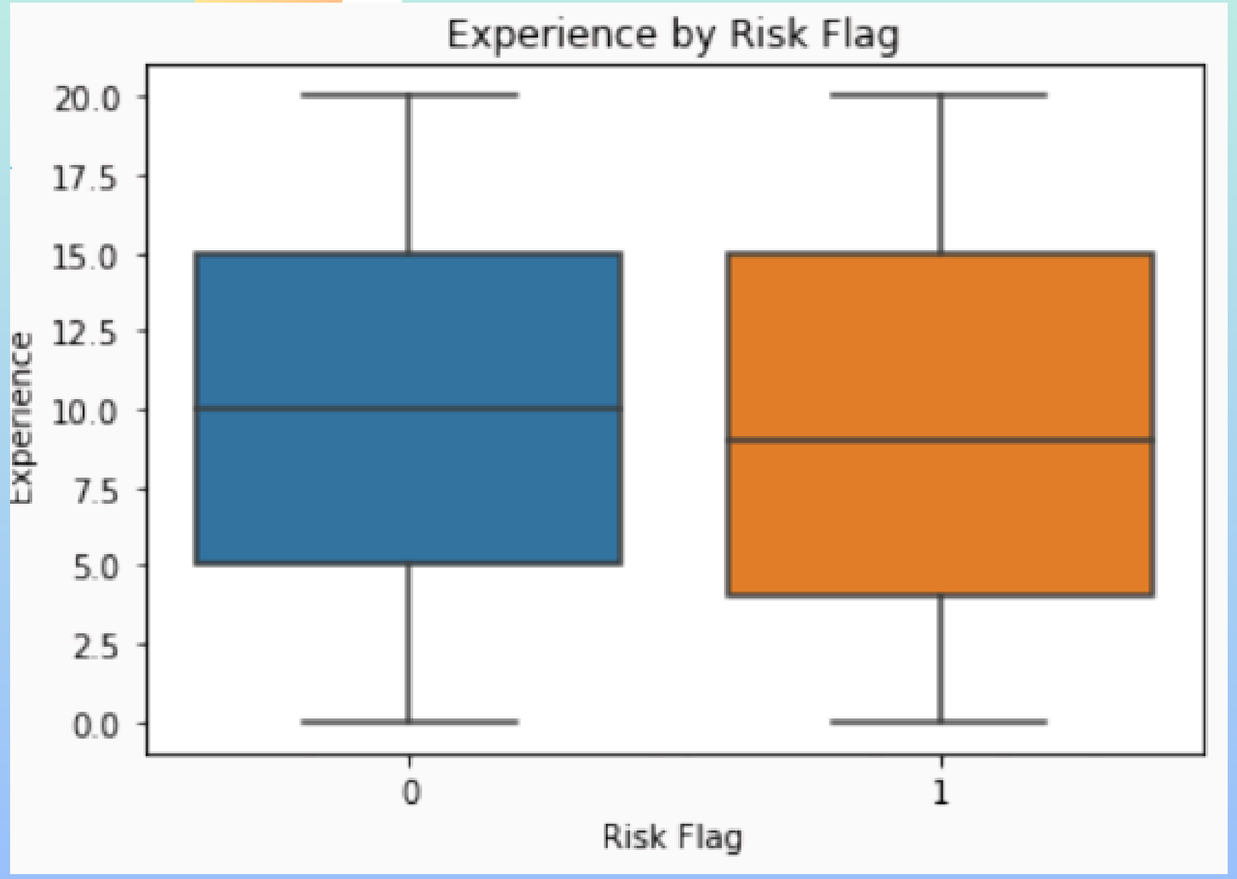
More than 200,000 of the applications get approved as they are flagged non_risky

Analysis Marital status vs Risk Flag .



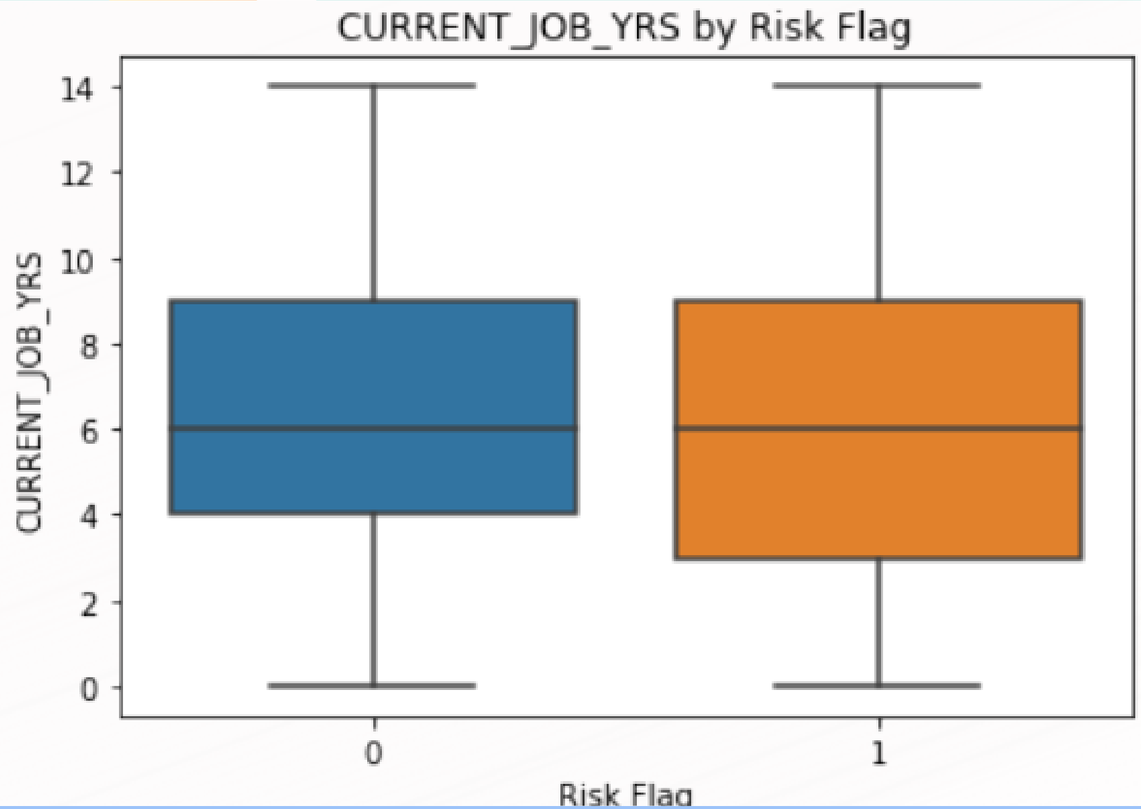
Most loan applicants are single and about 25,000 of the applicants are flagged as risky

Analysis of Years of Experience vs Risk Flag .



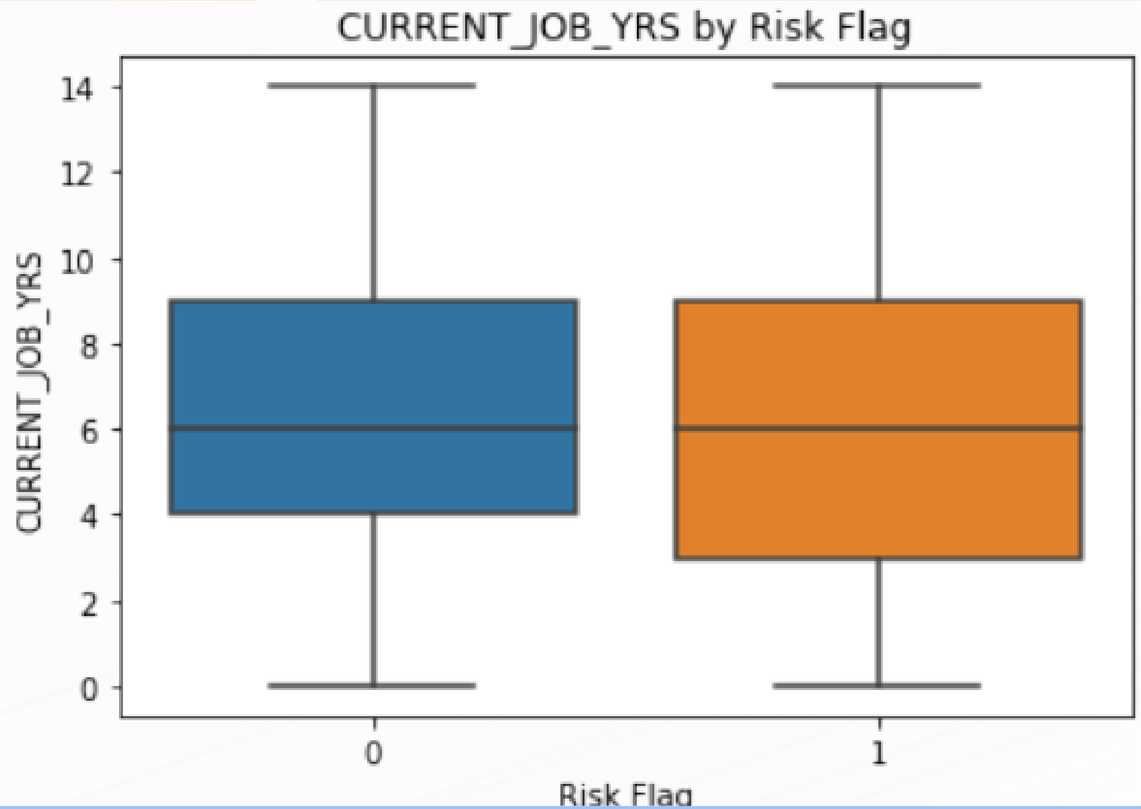
Applicants with more years of experience are deemed less risky

Analysis of Current Job vs Risk Flag



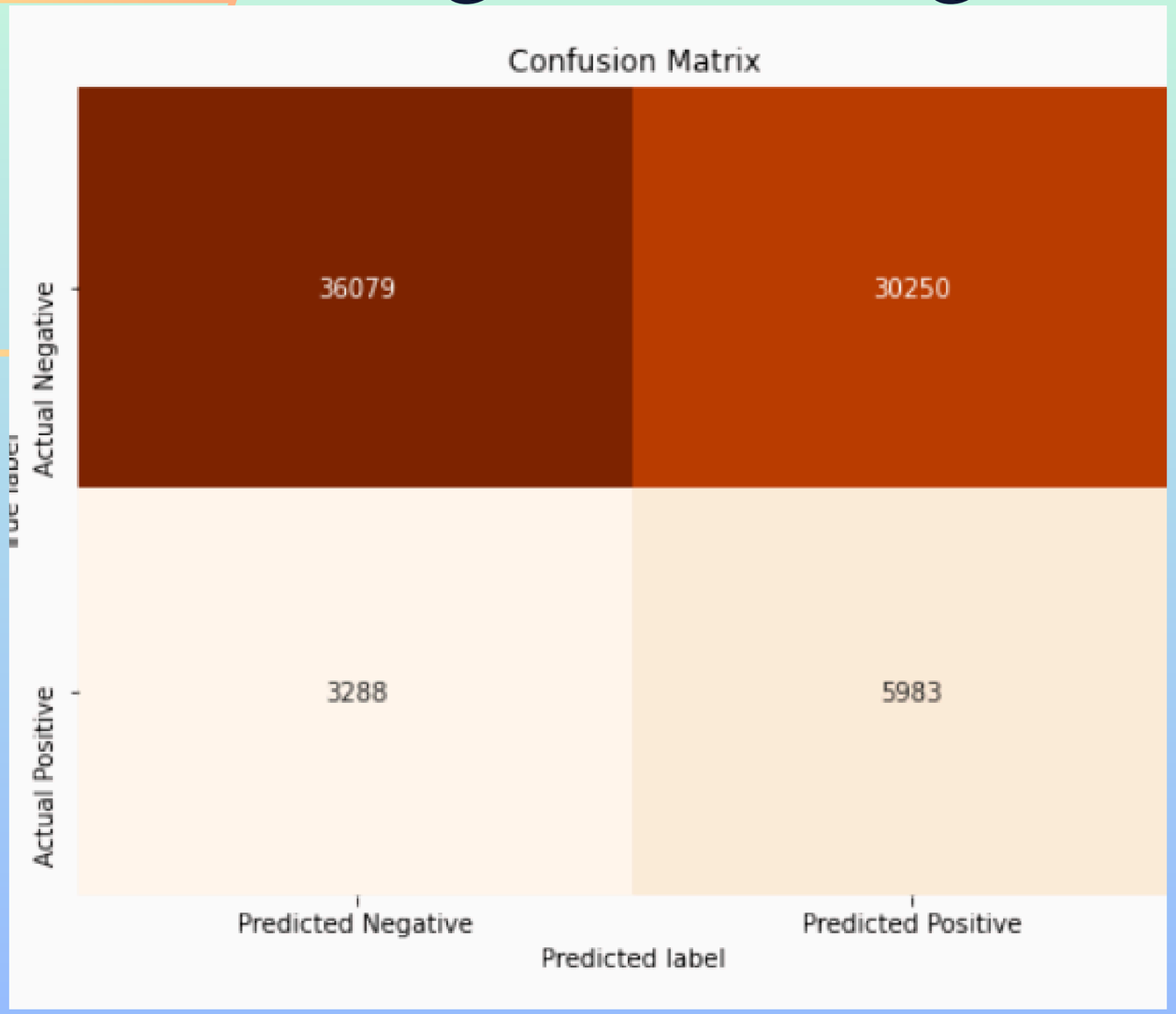
Applicants with short job tenures are flagged as risky applicants

Analysis of Current Job vs Risk Flag



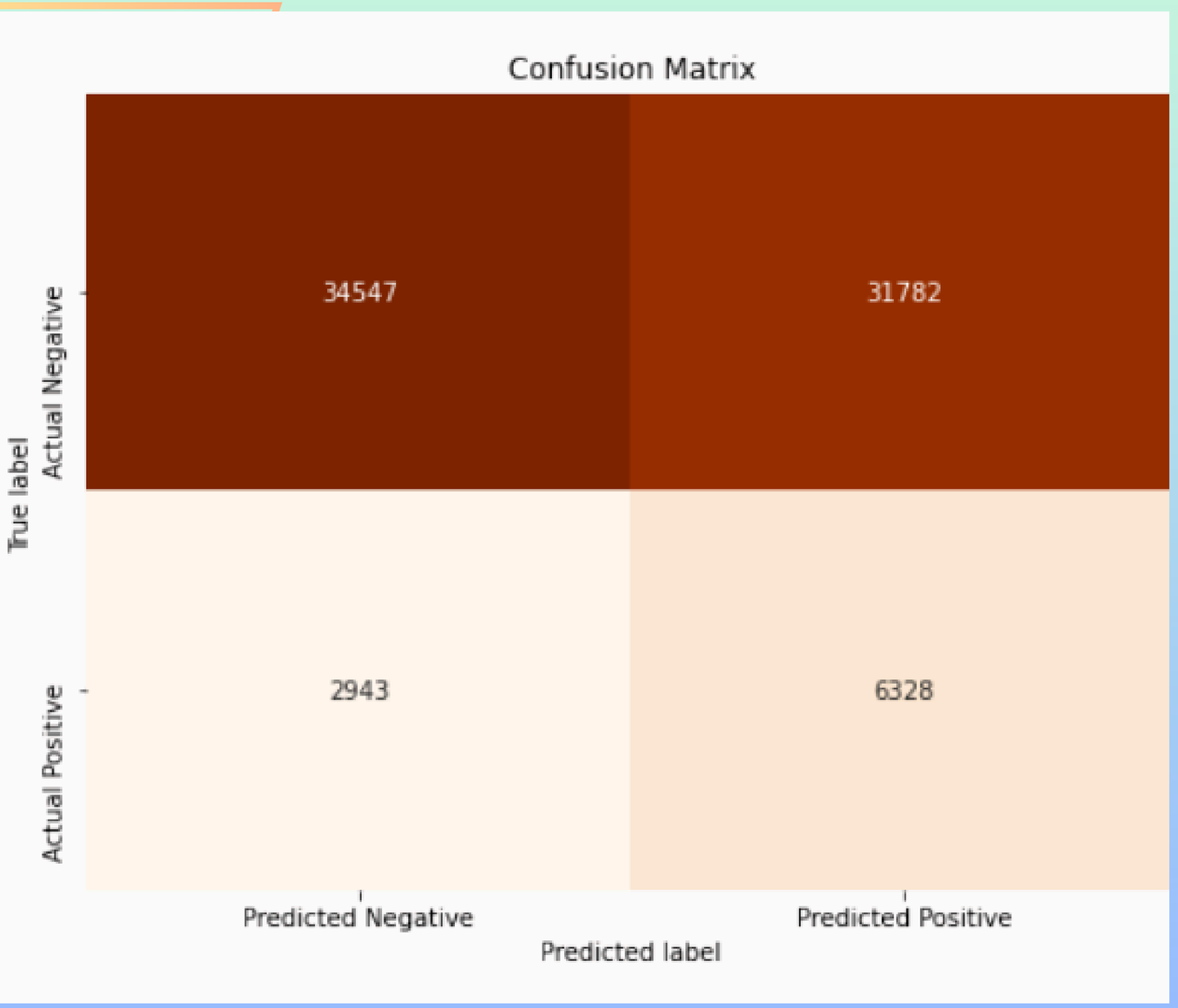
Applicants with short job tenures are flagged as risky applicants

Logistic Regression Model.



.Had an Accuracy of 55.67%
.Acted as our baseline mode

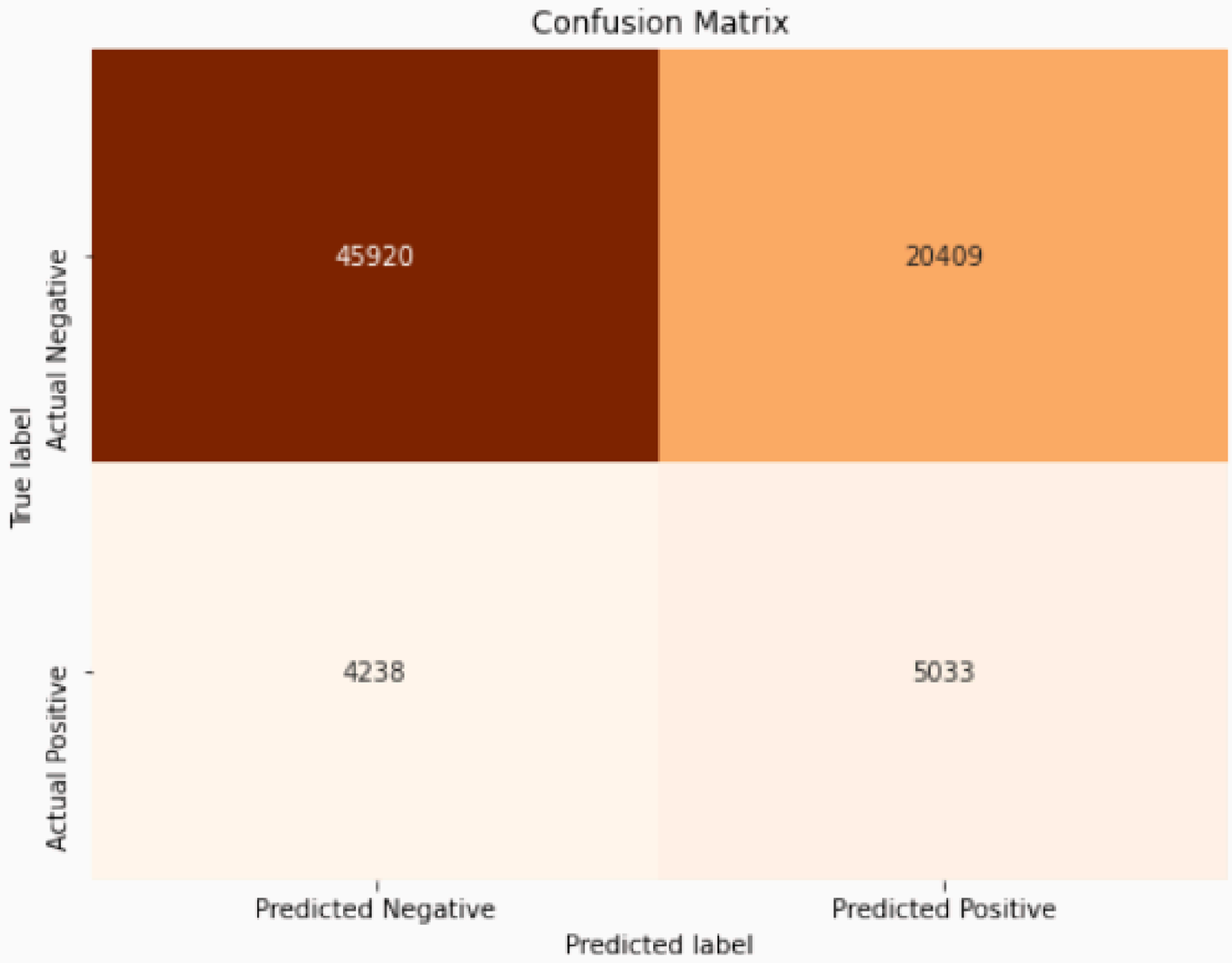
2nd Decision Tree Model.



.. Had an Accuracy of 54.1%

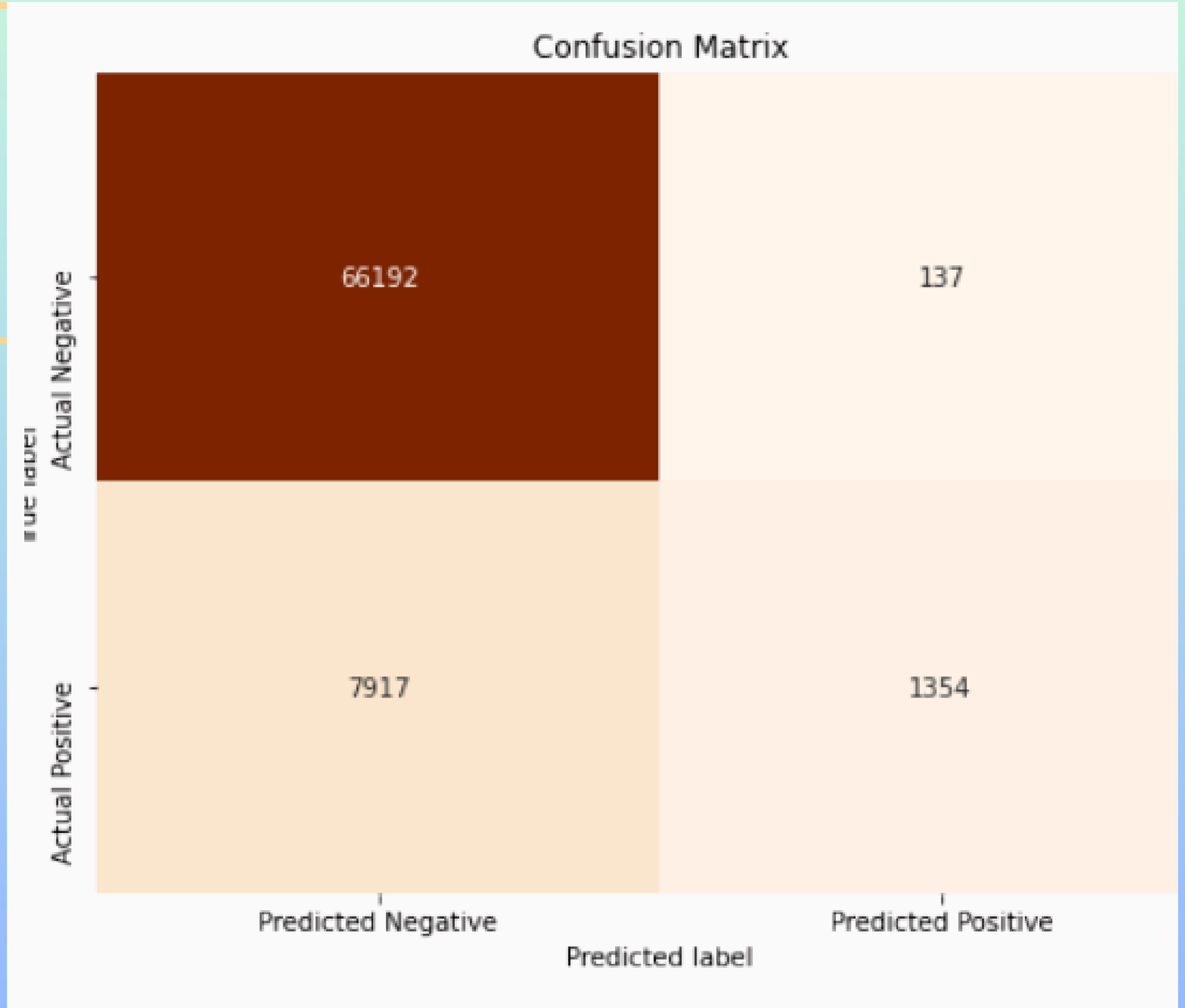
.. The model was not overfitting to the training dataset

Random Forest.



--Had an Accuracy of 68.9%
--The model detected 54.3% of all positive instances

XGboost.



--Had an Accuracy of 89.3%
-- In every
situation
it predicted a risky applicant and
was correct 90% of the time

Recommendations.

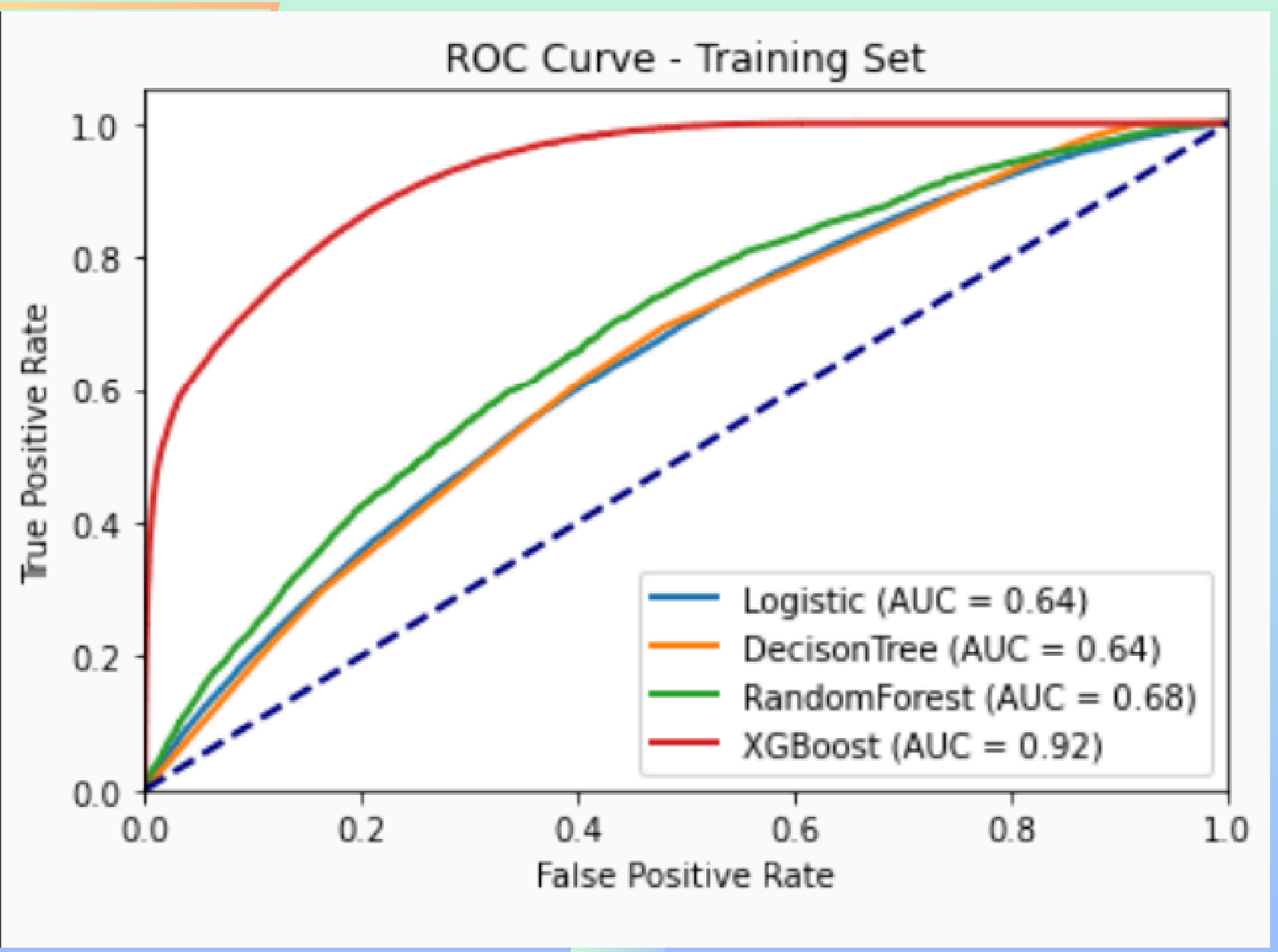
1. Provide training for stakeholders and implement a feedback mechanism to improve the model continuously

2. Conduct fairness analysis to prevent discrimination against any group and nsure the model complies with relevant regulations

3 Explore the use of advanced machine learning techniques like deep learning to further improve model accuracy and continuously explore new features that can improve the predictive power of the model.

4 Continuously monitor model performance and Plan for regular updates and retraining with new data to maintain accuracy and relevance.

ROC curve.



--The XG Boost model is closest to the topleft corner: Indicates a better performing model.
--The model is correctly identifying positives while minimizing false positives.
Followed by the RandomForset Model, DecisionTree and Logistic Regression model

Next course of Action

---Ensure that the model complies with relevant regulations and ethical guidelines, particularly those related to fair lending practices.

--Identify and create new features that could improve the model's predictive power

--Set up monitoring to track the model's performance over time and implement a maintenance schedule for regular updates and retraining

Thank You

0793923656

martinkandie09@gmail.com

Moringa School 1st Floor, Ngong Lane Plaza, Ngong Rd, Nairobi, 00100, Kenya