



Prediction If A Bank Customer Would Subscribe A Term Deposit

Analysis by Aykut Avci

Table of Contents

Introduction & Objective

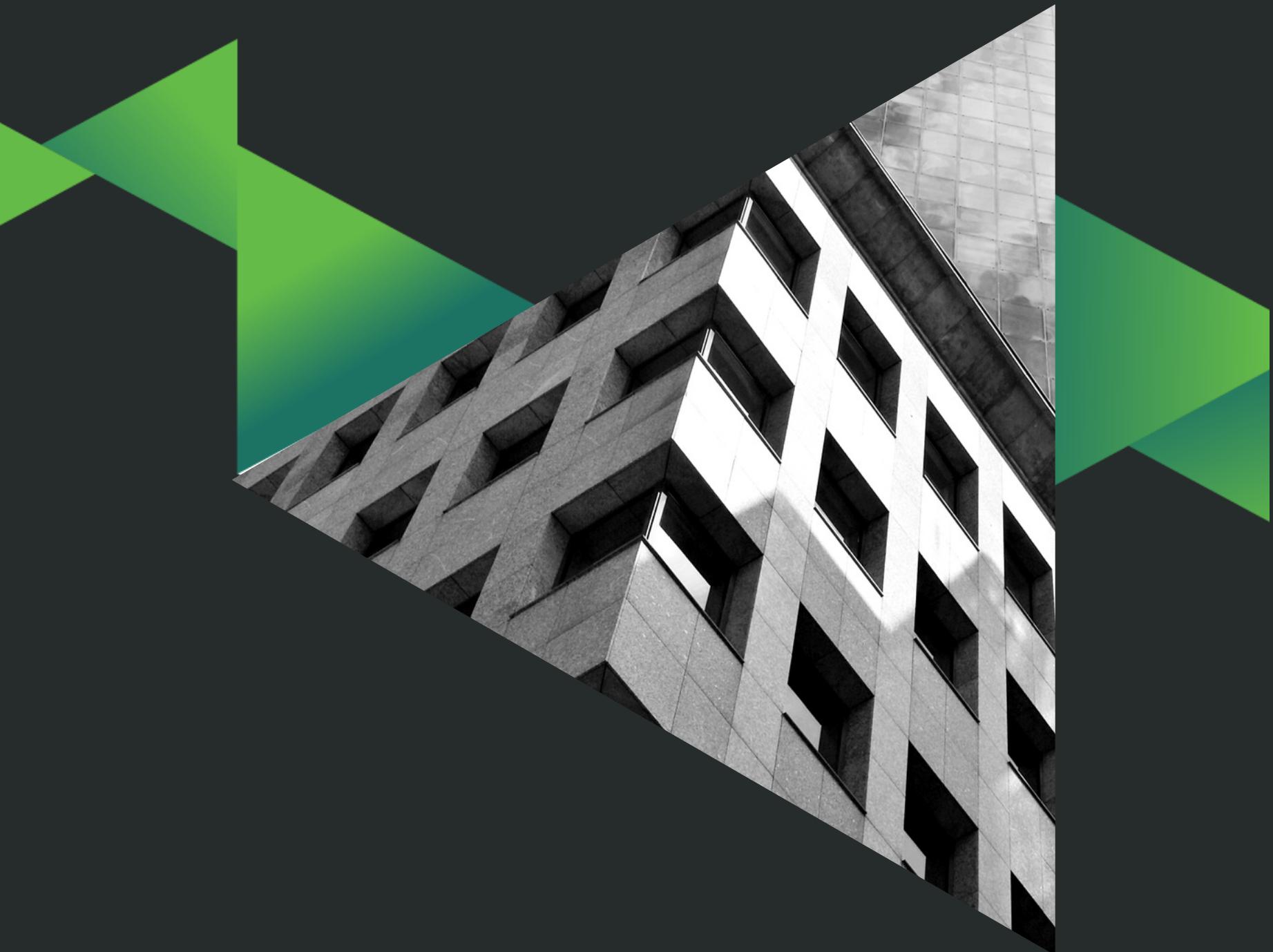
Data Overview

E.D.A

Model Presentation

Conclusion





Introduction

- Customer, Marketing and Bank Related Features

Purpose

- Predicting the customers who are likely to accept marketing campaign (locking up money in the account for a period of time)
- Reducing Costs

Dataset



SOURCE:

- UCI Machine Learning Repository

RAW DATA:

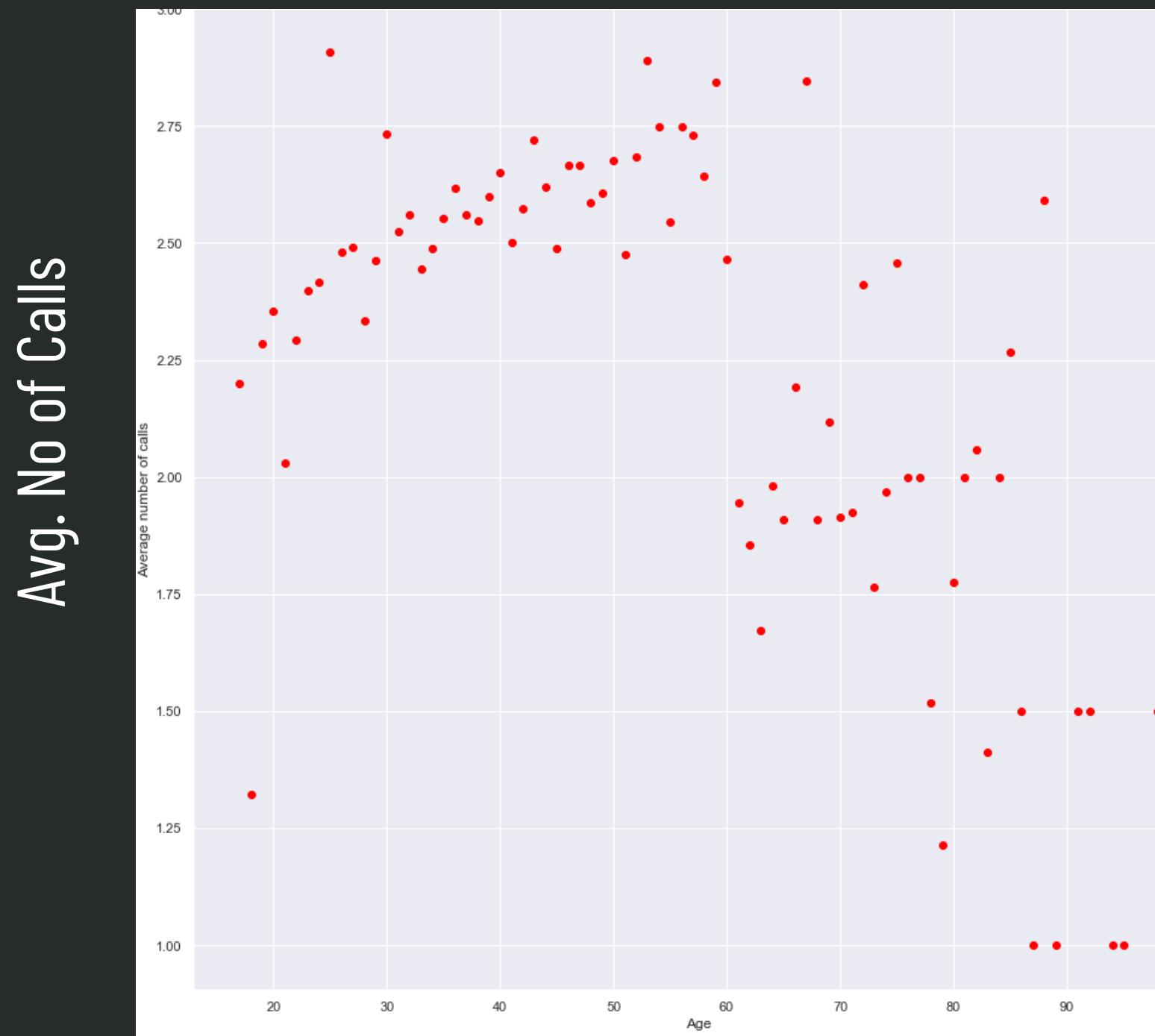
- 20 Input Features & 1 Target Feature
- ~ 41k Records

CLEAN DATA:

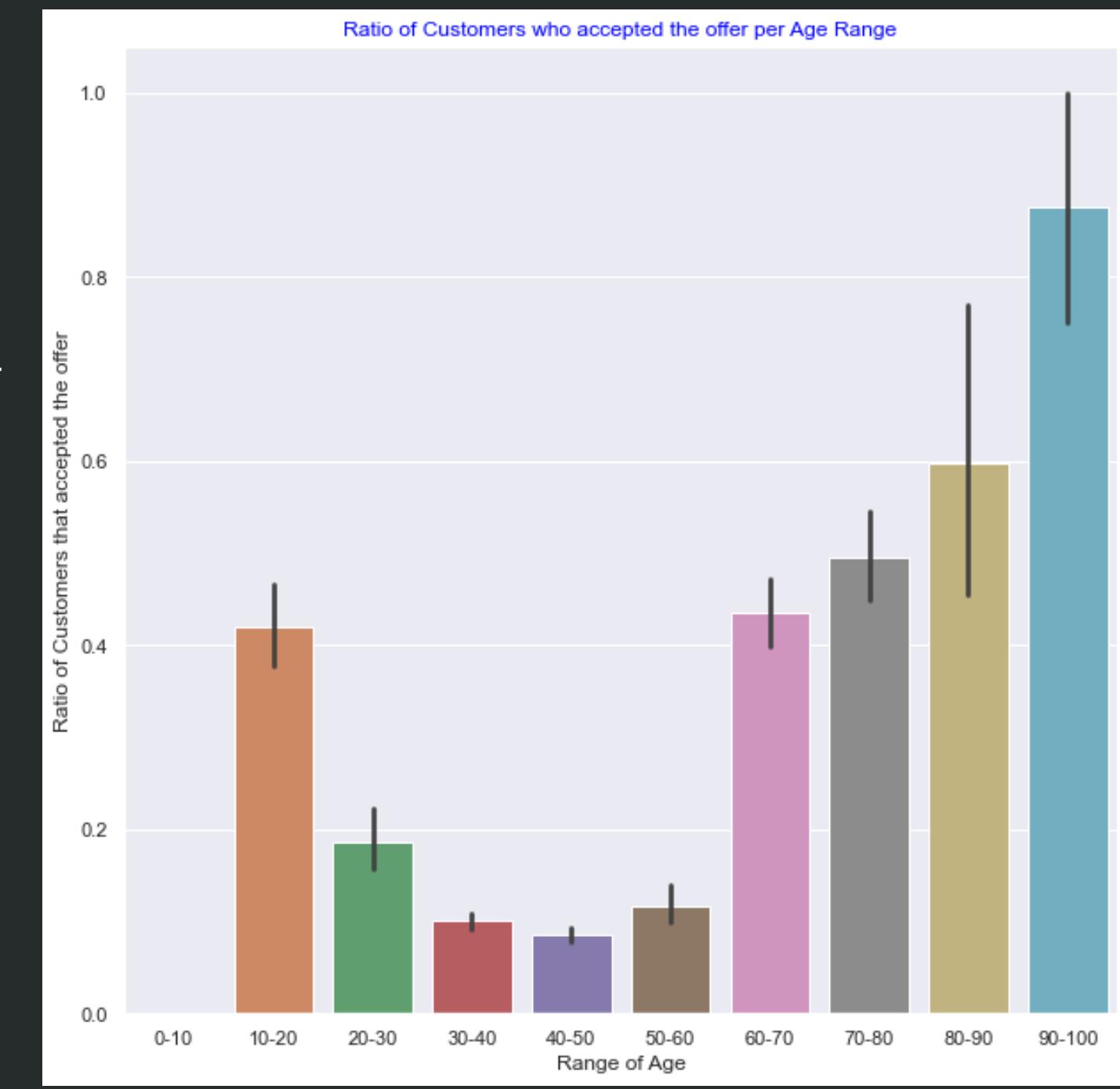
- 18 Features
- ~39k Samples

E.D.A.

Effect of Age in Customer Behaviour



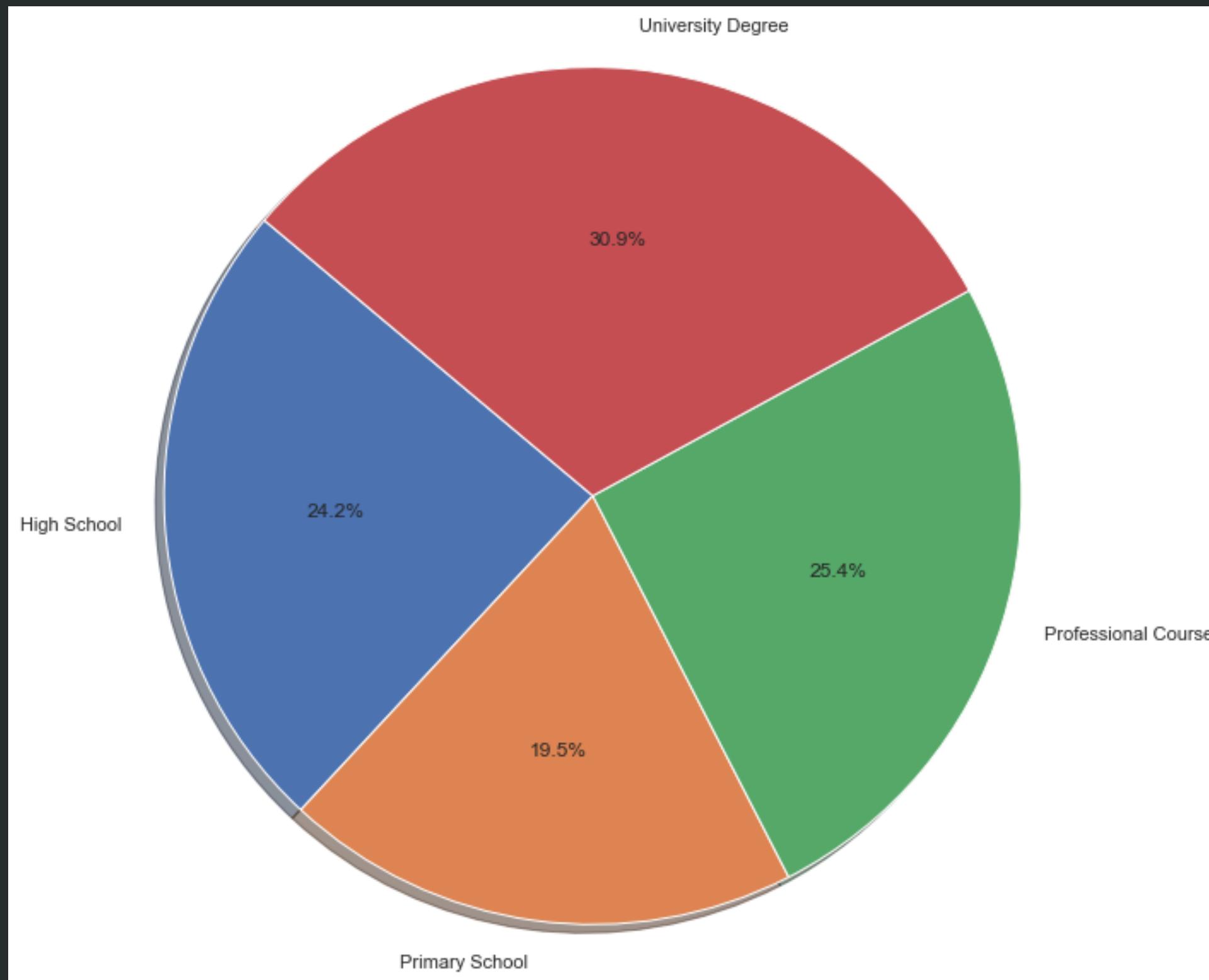
AGE



AGE RANGE

E.D.A.

Effect of Education in Customer Behaviour



Data Processing

Deleting Useless Features & Samples

- lc-duration
- c- default

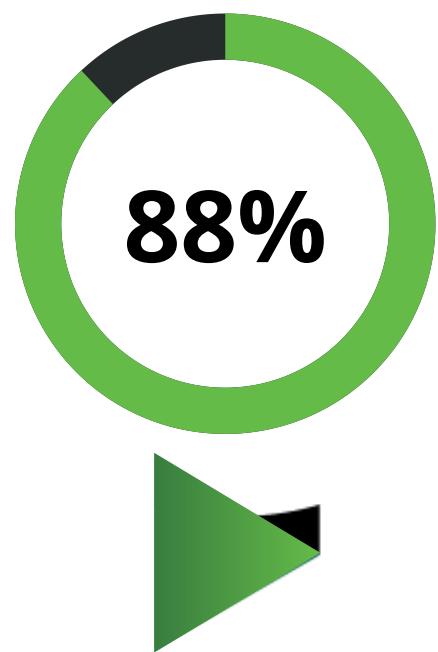
Replacing Unknown Values

- Unknown entries in categorical features
- Replaced with the mod of the column

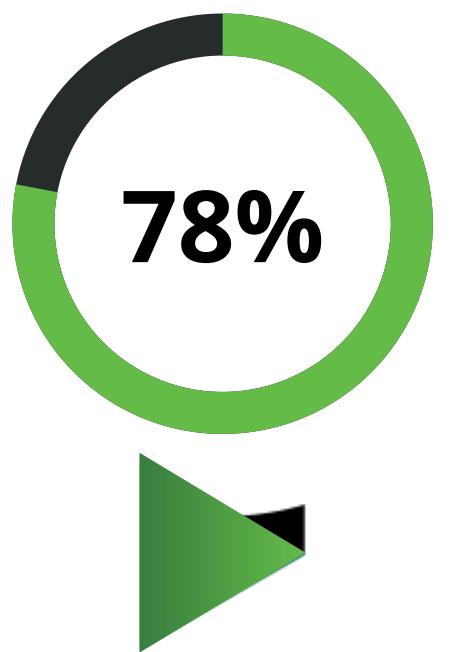
Handling Imbalance in Target Feature

- ~ 35k Negative & 5k Positive Output
- Upsampled using SMOTE

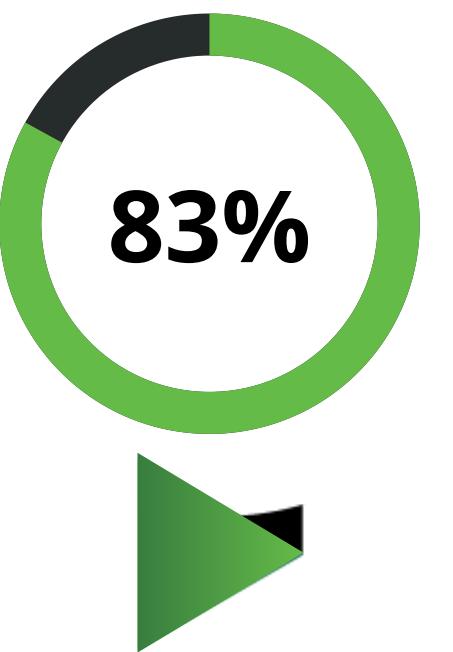
Model Comparison



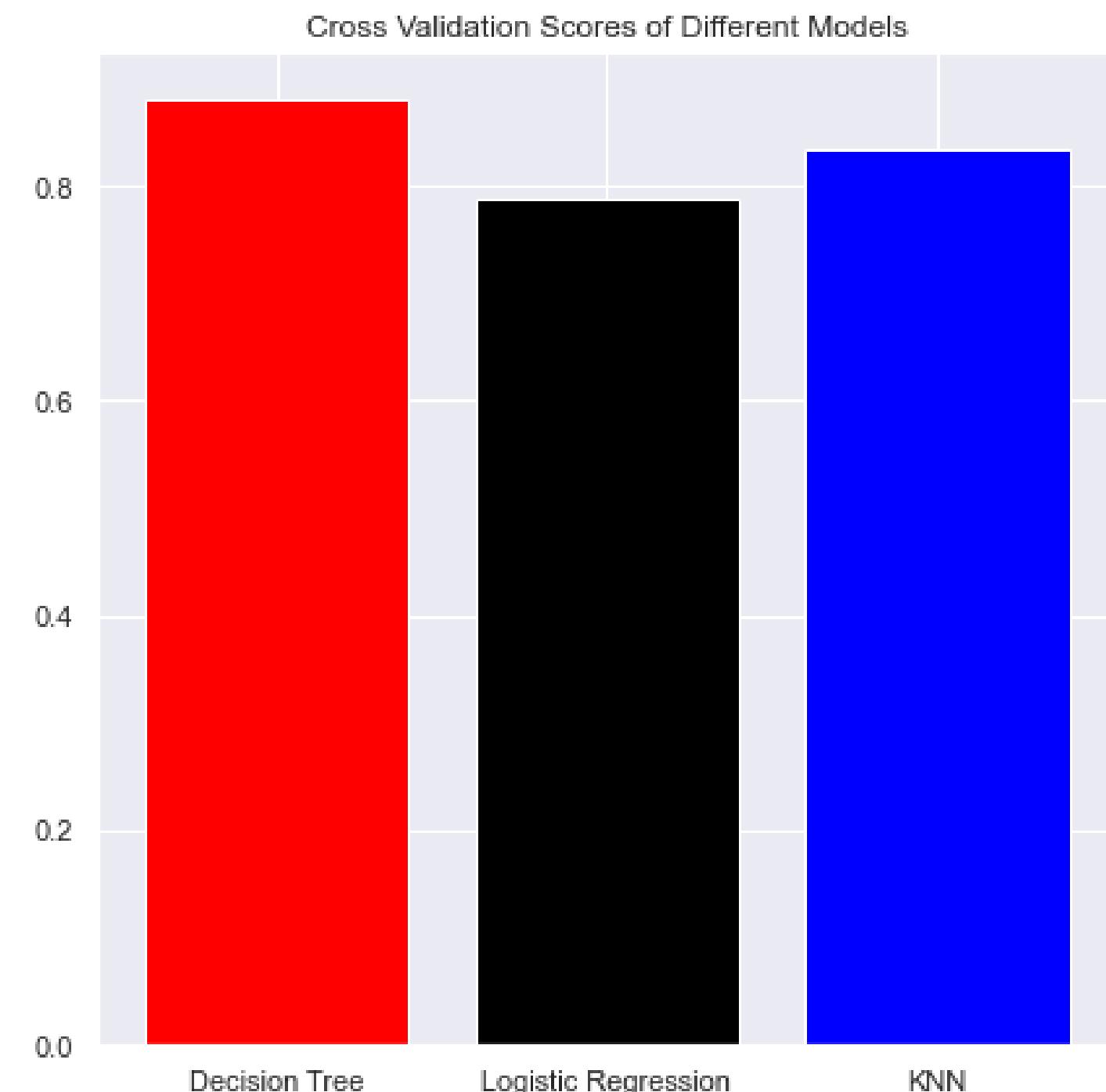
Decision Tree Classifier



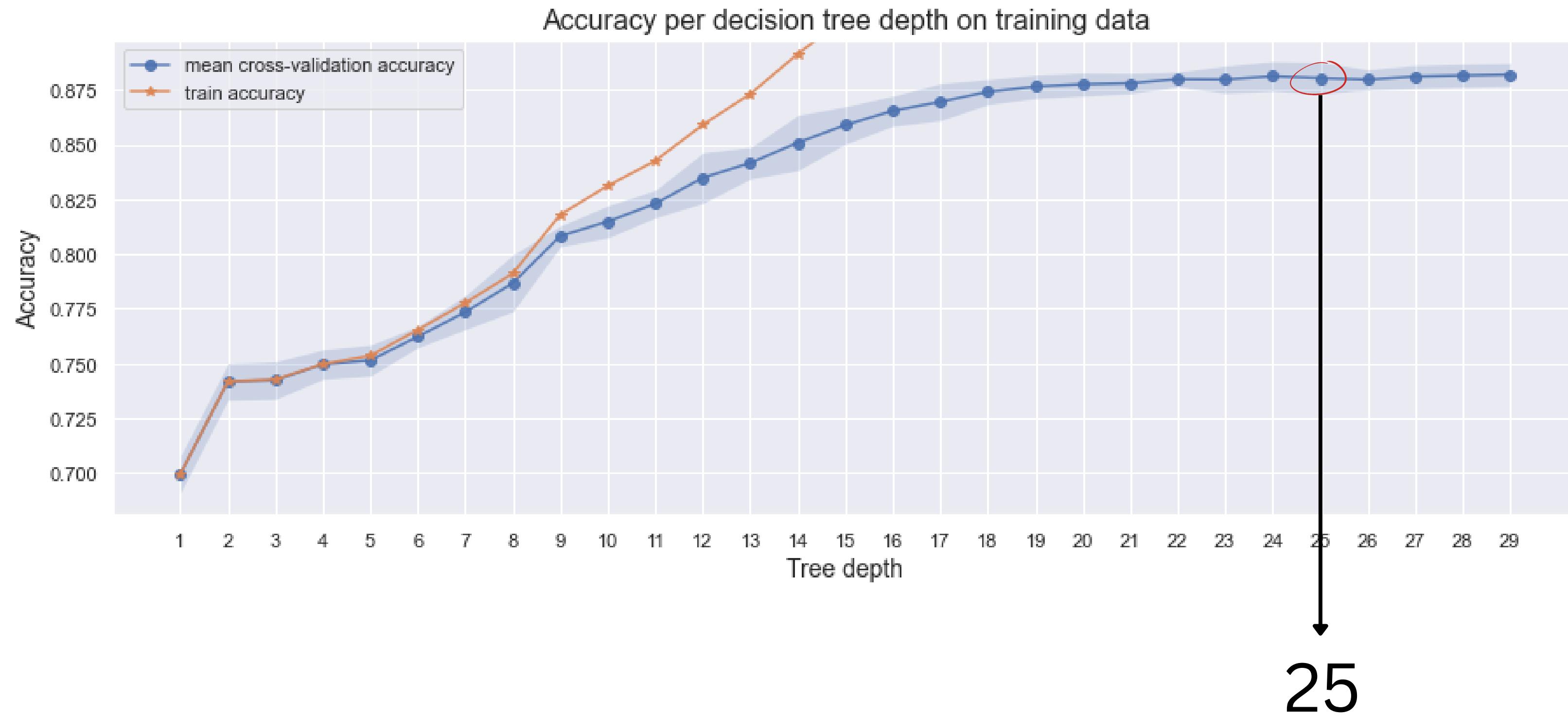
Logistic Regression



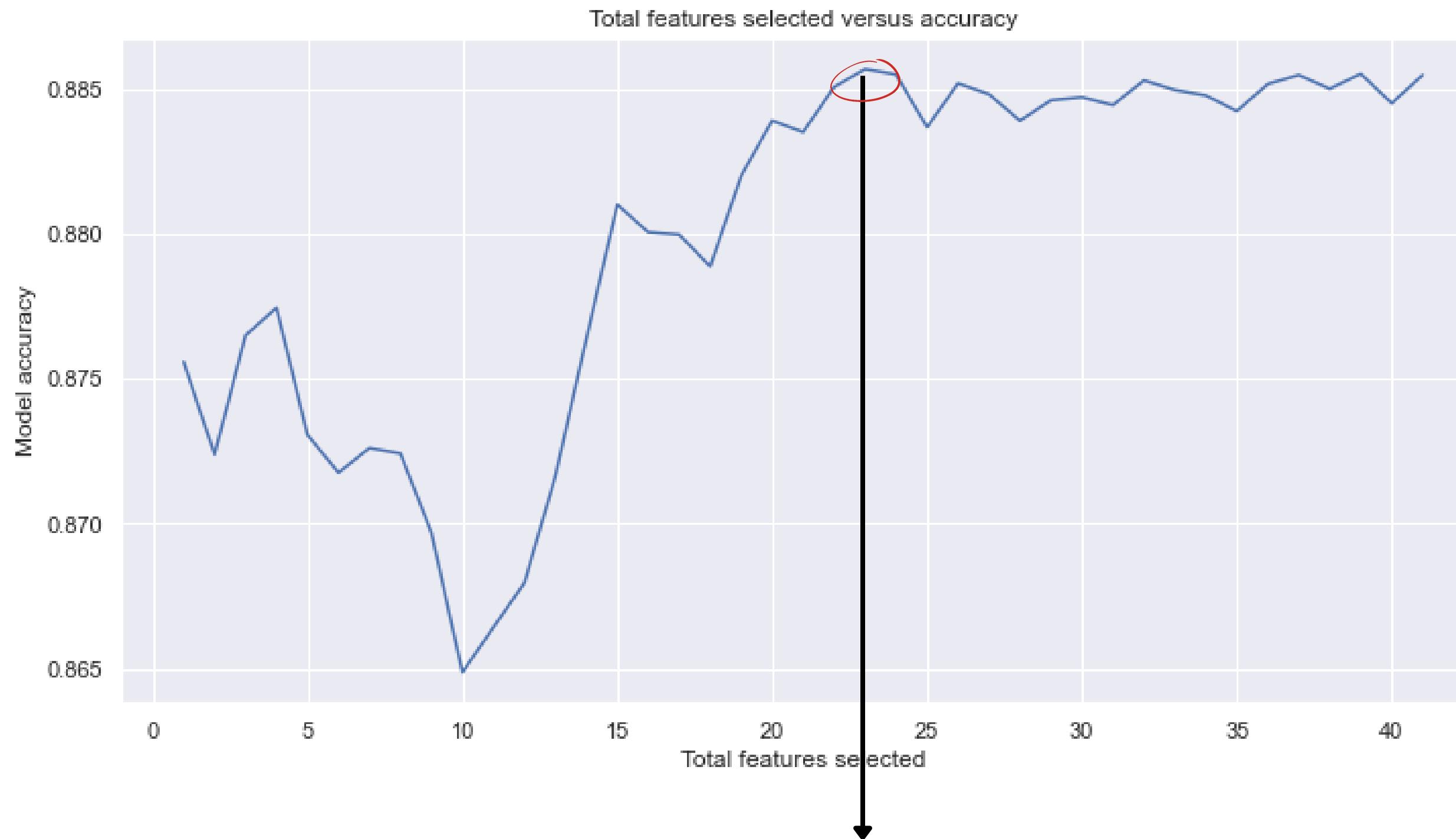
K-Nearest Neighbours
Classifier



Choosing Best Decision Tree Depth



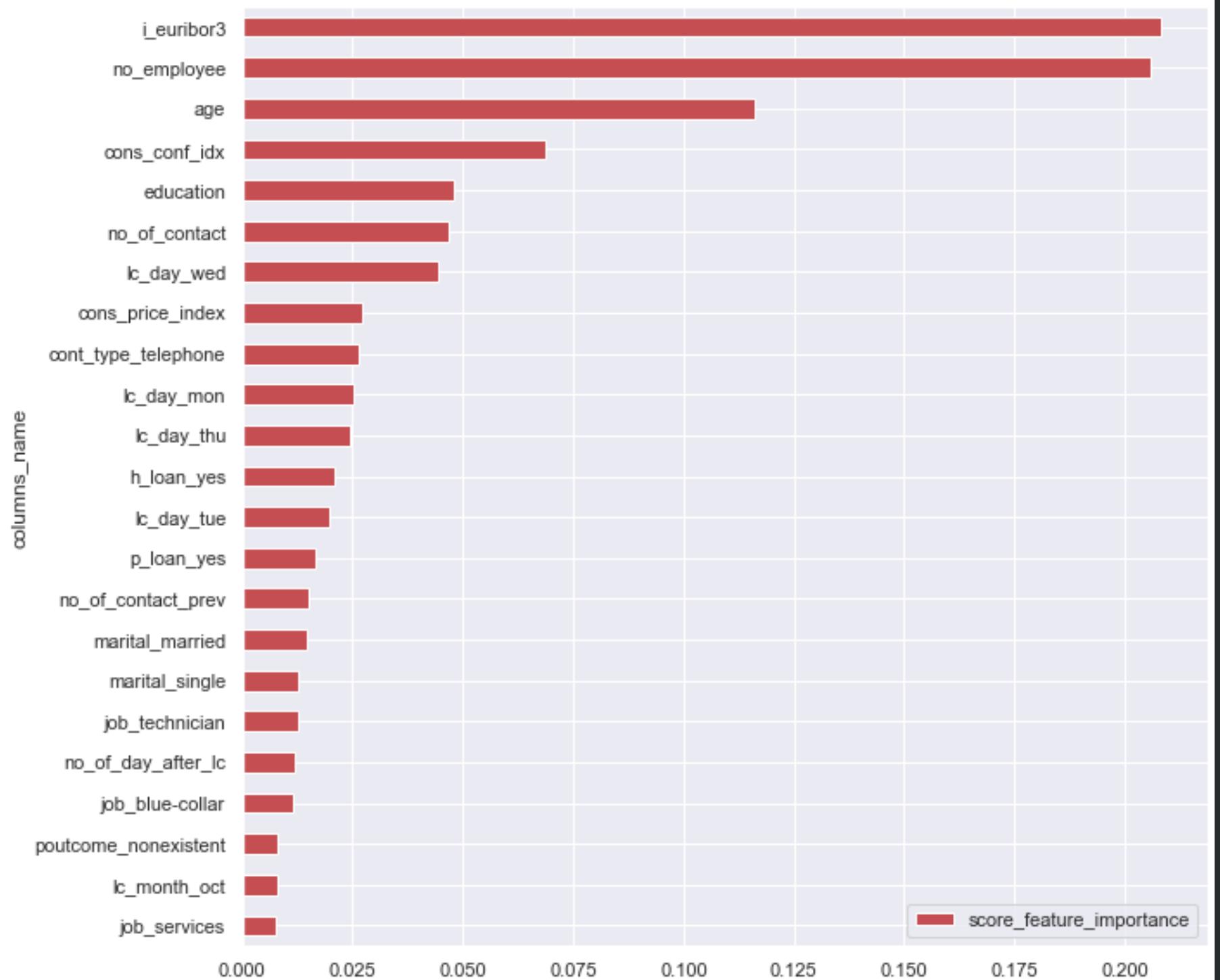
Elimination of Statistically Insignificant Features



23 Features

Feature Importance

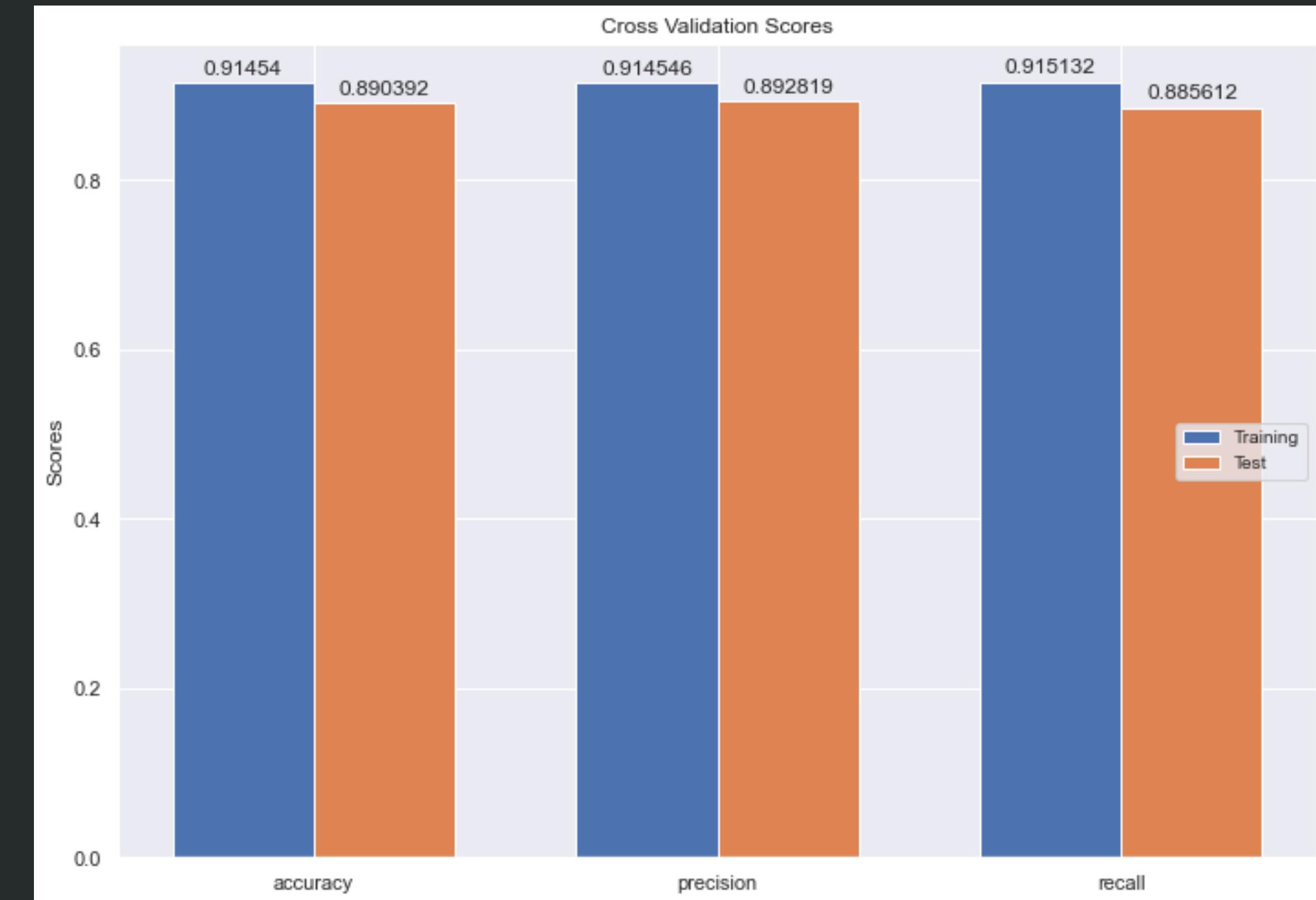
- Age
- Education
- Number of Contact





Random Forest with
Hyper Parameter Tuning

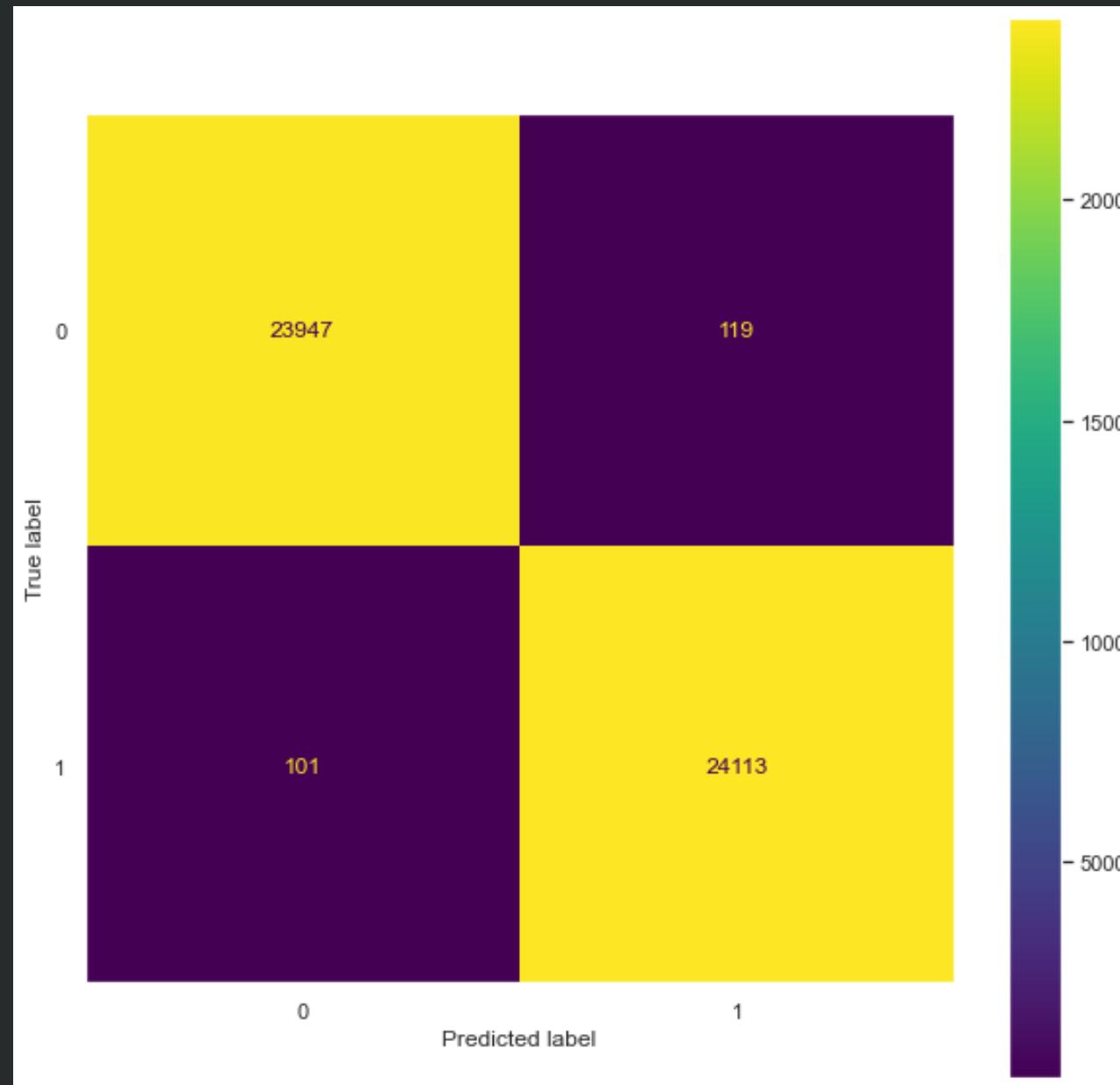
Scores of Error Metrics with Cross Validation



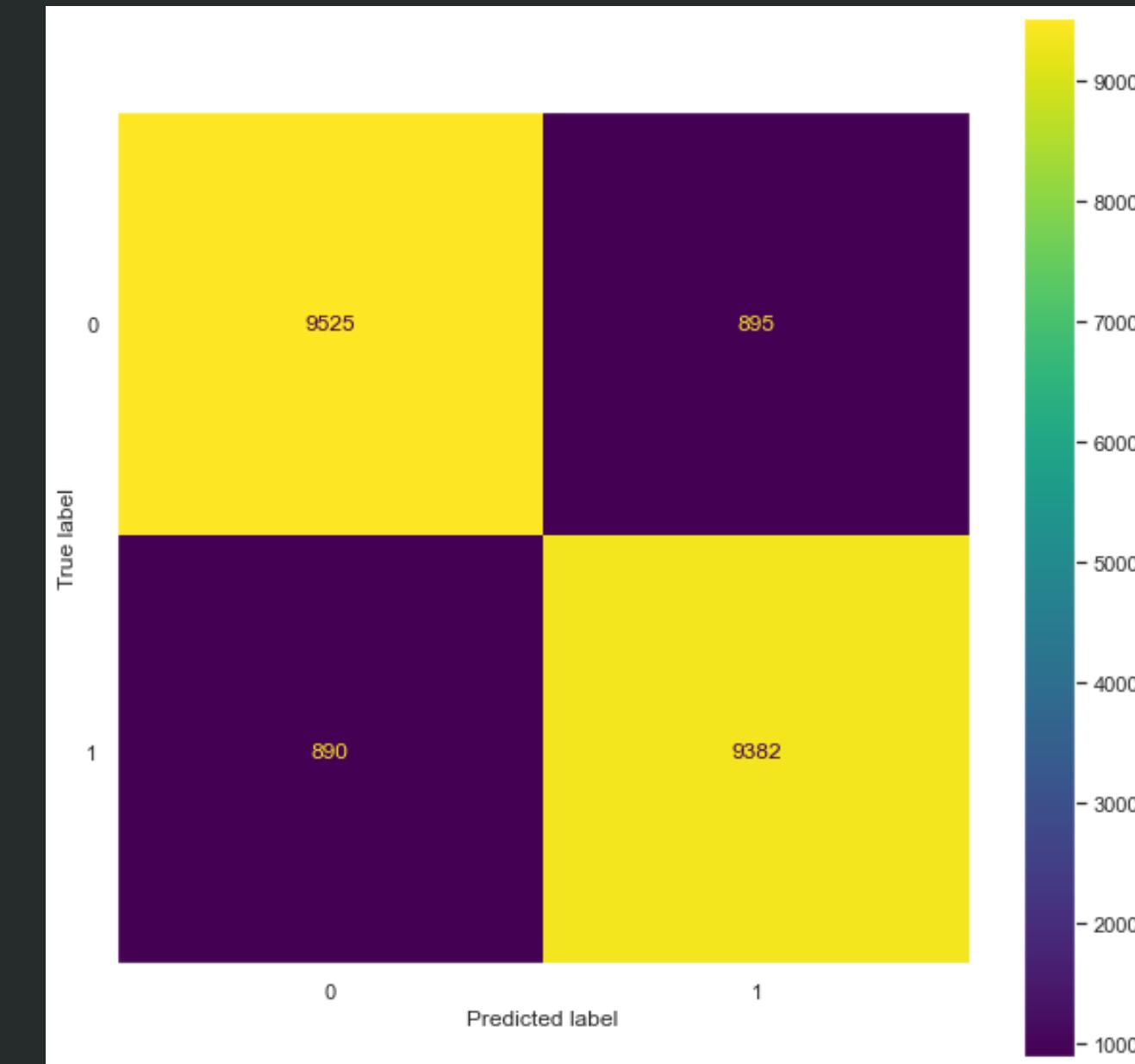
91% and 88% Recall Score for Training and Test Set Respectively

Confusion Matrices

Training



Test



Conclusion

- 89% Accuracy, 88% Recall on Test Set
- Age has the highest weight in customer features
- Education is directly correlated with offer acceptance
- Predictive power can be improved with more balanced data
- Customer segmentation based on age and education



THANK YOU