10Alytics Project

Predicting Bank Product Uptake

machine learning

VICTOR ETIM

Problem Statement

01 We want to know how many customers will subscribe to the product.

02 There is a need to target the right kind of customers thereby minimizing campaign expenditure.

Project Objective

01 Build machine learning model which will determine whether customers will subscribe to the term deposit.

02 Determine which model is most effective for the campaign's prediction.

Because this is a marketing campaign and the number of true positives are more important than false negatives, the prediction confusion metric is of more importance than a recall metric.

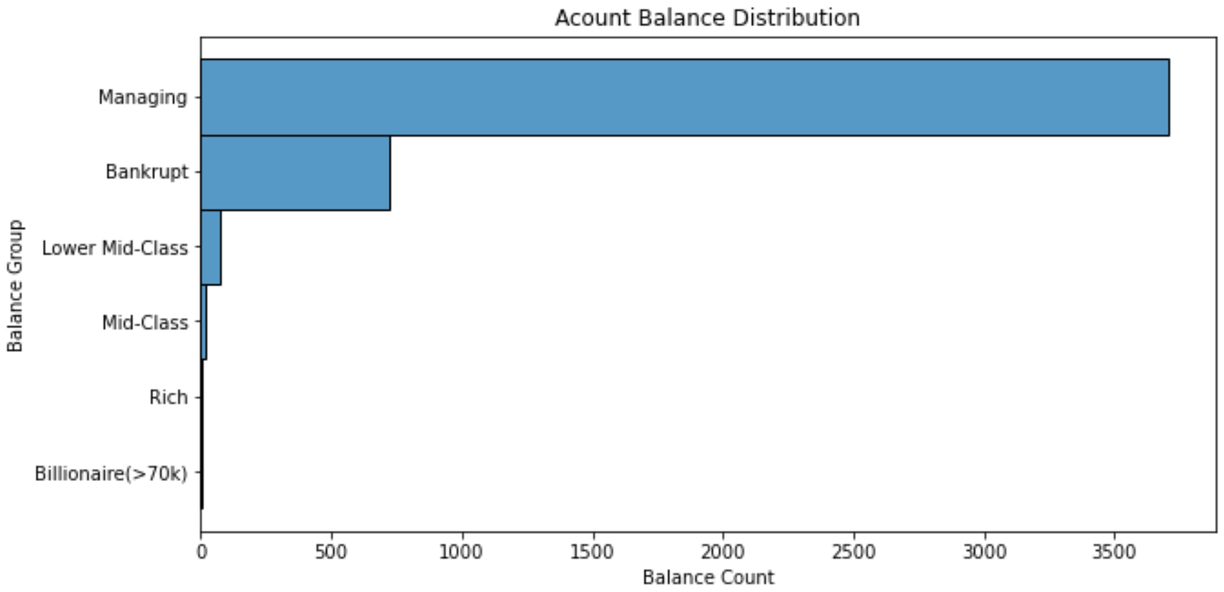
VISUALIZATION

Insights from Visualization

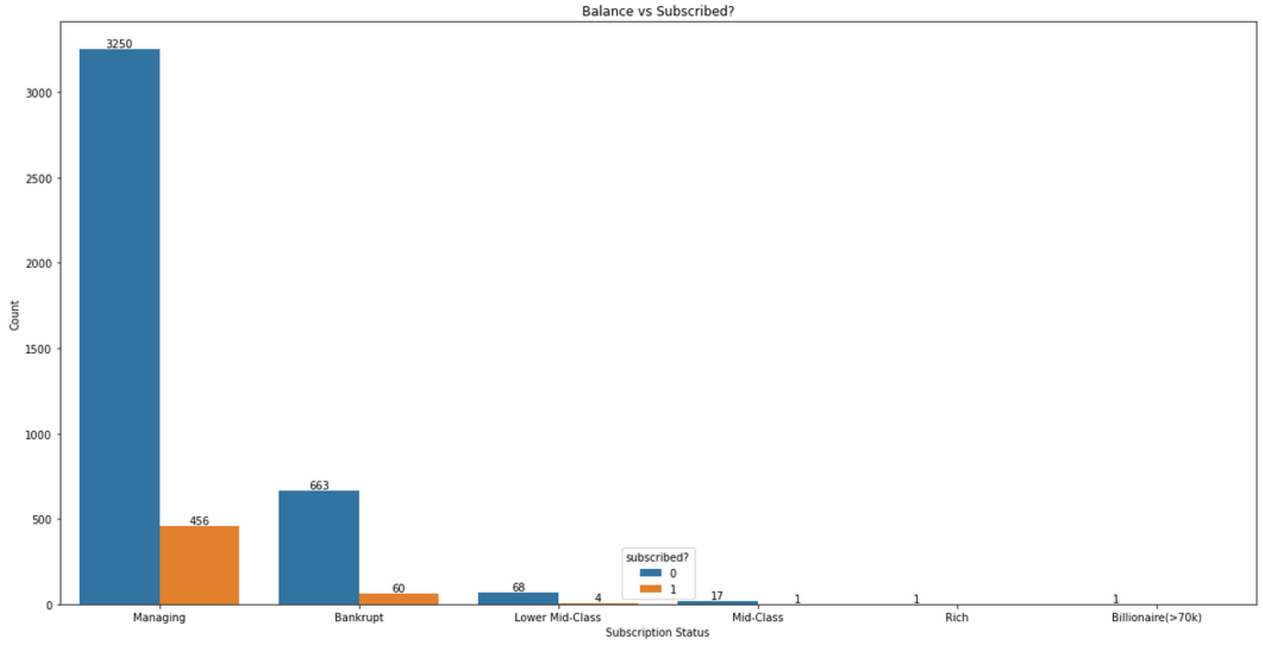
By observation, most of the people who obviously subscribed have a positive account balance.
From this bivariate visual, we can see and glean a number of things:

- most people are normally not interested in these term deposits,
- The people in the Managing group (<=10k balance) are the ones most likely to subscribe.
- the more money people have in their balance, they are less likely going to do a fixed deposit; maybe because they would rather choose to be investors.

Account Balance Distribution of Bank Customers



Plot of Account Balance against Subscription Result



Insights Summary

Here are a couple of things I discovered,

01 The preferred metric for forecasting the product uptake campaign success is Precision.

02 Most of the feedback data on the campaign outcome were reported as Unknown.

03 Majority of the customers have an account balance less than 10,000.

04 Customers with amounts saved in their account over 10,000 are less likely going to patronize the product.

Action Point

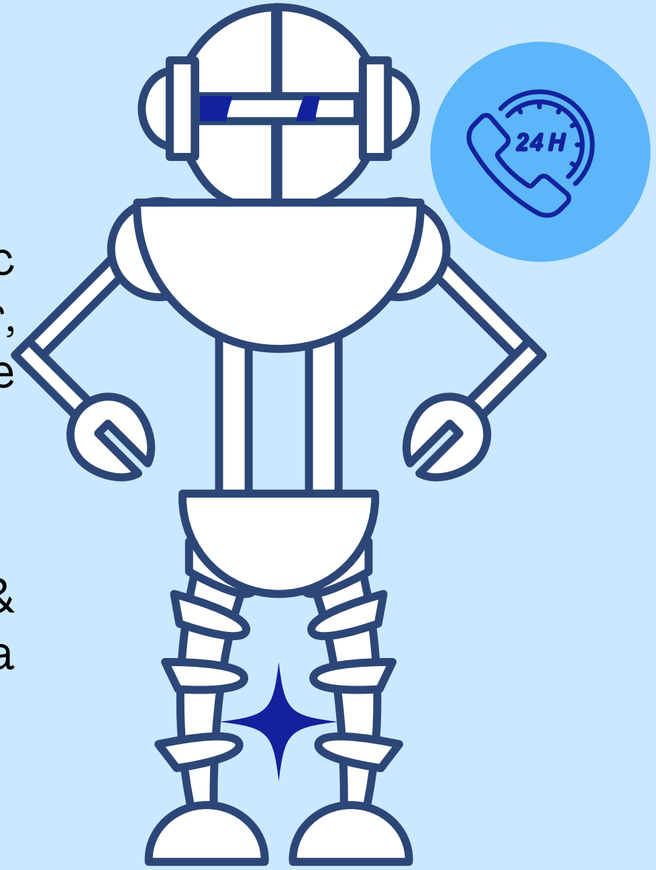
Target People with amounts <= 10k in their account balance.

Model Evaluation

Initial Modelling

- Model options were Logistic Regression, KNeighbors Classifier, Random Forest Classifier and the Support Vector Machines (SVM).

- Initially, both Random Forest & SVM look the best option with a tied prediction of 68% each.



Optimization Choice

- In order to carry out optimization, standardization comes first.

- After optimization, SVM offers the best prediction potential of 76%.

Thank You