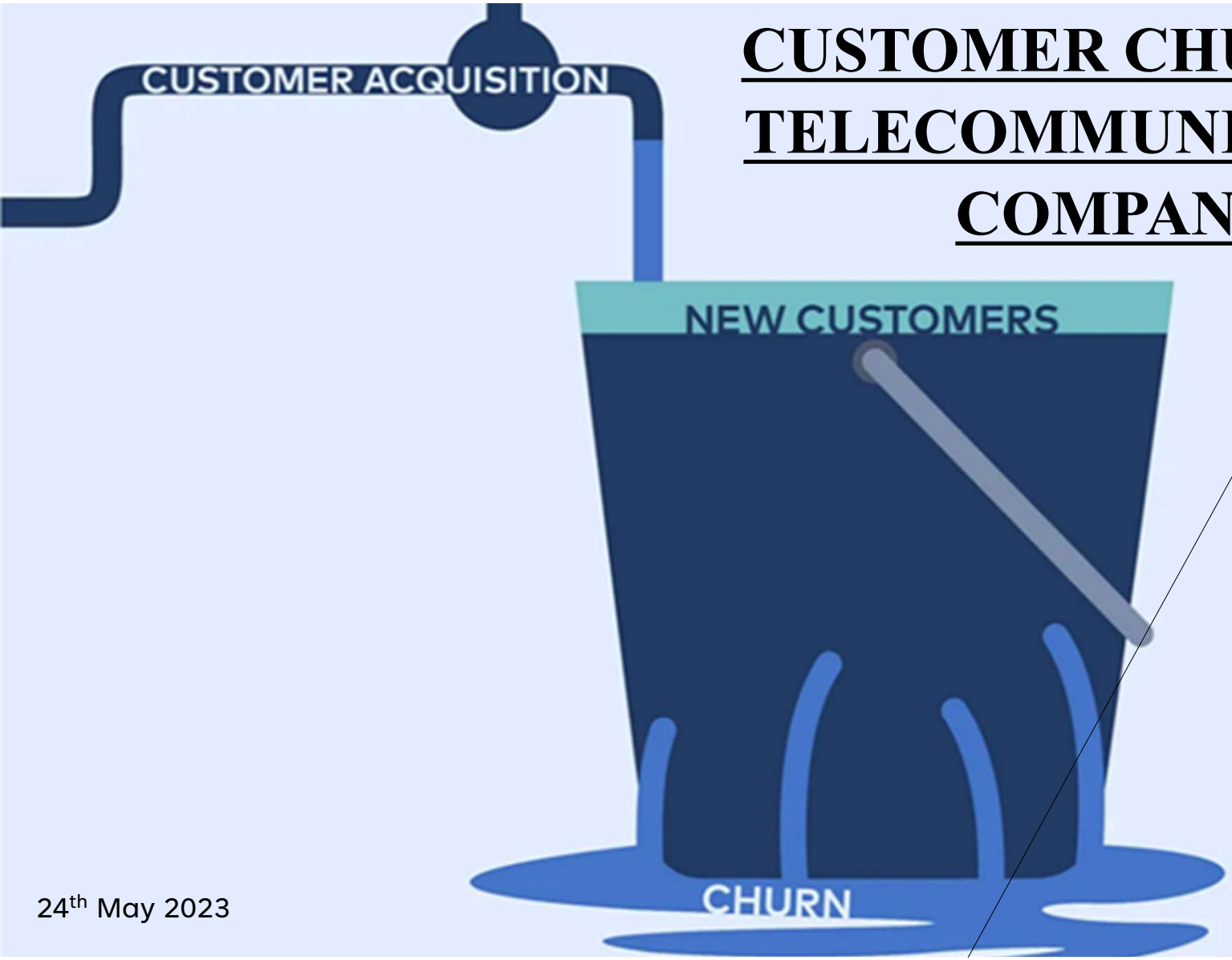
CUSTOMER CHURN IN A TELECOMMUNICATION COMPANY



24th May 2023

INTRODUCTION

This project seeks to build a classifier to predict whether a customer will soon stop doing business with SyriaTel, which is a telecommunications company.

It seeks to identify any predictable patterns among customers who have already left and use these features to predict customers who are likely to leave in the future.

BUSINESS PROBLEM

- **SyriaTel**, a telecommunication company is facing the problem of **an increase in the number of customers who leave the company.** A consistently high churn rate could result in the company quickly becoming **unsustainable.**
- **Attracting new customers** as a strategy is **not enough** to sustain the company for very long.
- It is therefore **important for the company** to increase the number of loyal and devoted customers by identifying the pain points across a company journey and accurately predicting the customers who are likely to churn and therefore targeting them using aggressive strategies to reduce these points of friction

OBJECTIVES

- To **build a machine learning model** that can accurately predict customers who will churn based on the features in the dataset.
- To **rank features these features** in the dataset according to their order of significance

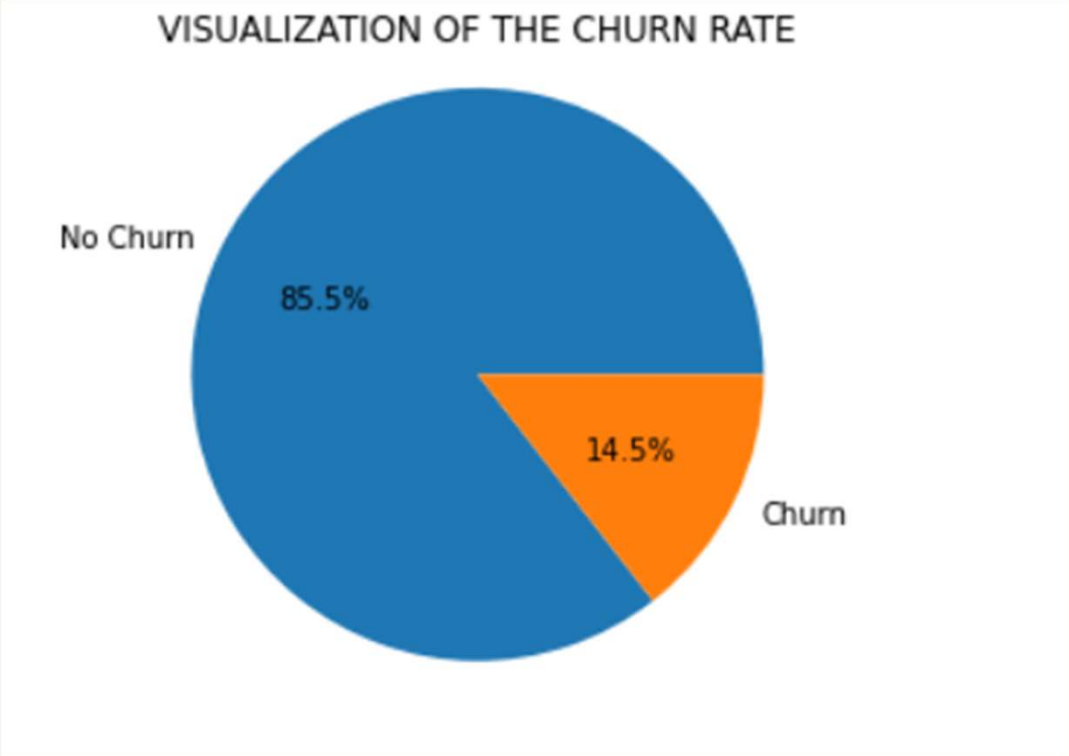
DATA UNDERSTANDING

DATASETS USED

- The dataset used is from SyriaTel Telecommunication company.
- Each row represents a customer, and the columns contain customer's attributes

DATA ANALYSIS

A plot showing the distribution of churn and no churn from the company

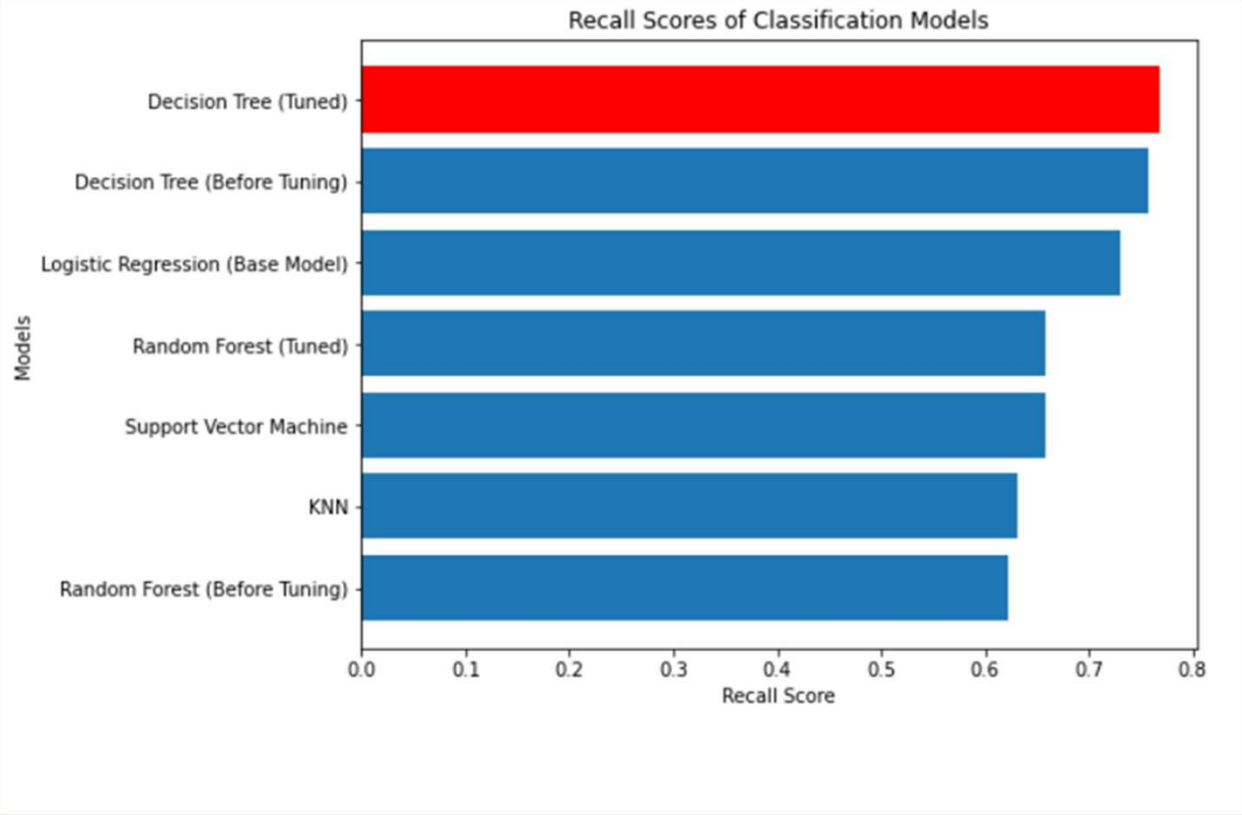


Analysis of the target variable:

- From the chart shown, there are more samples for customers without churn than for customers with churn.

MODELING

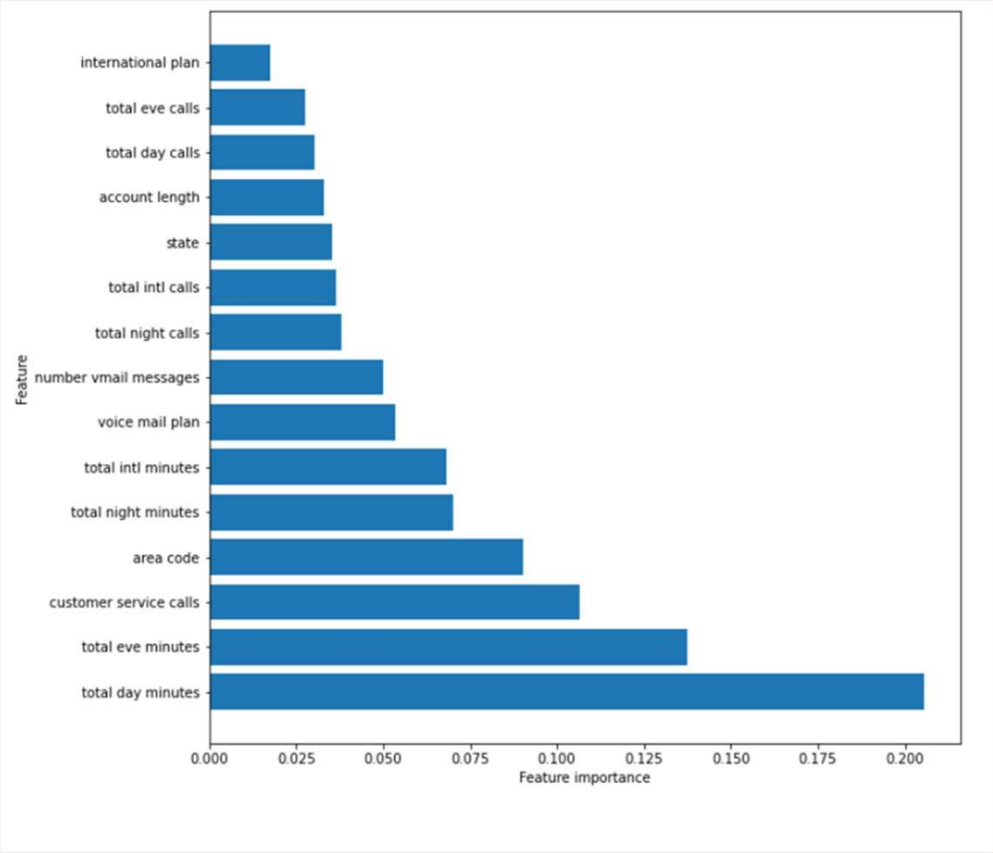
A plot showing the performance of all the models used



- The best performing model is the **Decision Tree Model**
- The worst performing model is the untuned **Random forest model**
- **A weakness** of the model is that due to the metric that has been optimized, it might have a **higher number of false predictions for churn than is true**
- However, **this is acceptable** as we would rather predict that a customer will churn and they fail to churn than fail to predict churn and they do.

EVALUATION

1. A plot showing the most significant and least significant features



- The most significant features were the total number of day and evening minutes spent on a call.
- The least significant features were the number of calls and whether or not a customer had an evening plan or not.
- This shows that it is the customers who make **long calls** and **not those who make many calls**, who are most likely to churn

CONCLUSION

1. The **model that meets the success criteria**, in predicting customers who will churn, of a sensitivity/ Recall score of 0.75 is the **Tuned decision tree model**.

2. The **most important feature** in predicting the churn rate is the **number of minutes** a customer spends on a call coupled with the **number of customer service calls.**

RECOMMENDATIONS

- SyriaTel should ensure effective customer service in order to meet customer's expectations and analyze customer interactions
- The company should also look into the call charges in comparison to their competitors, and model their charges competitively

NEXT STEPS

The next steps in this project would be to investigate how the area code ranks so highly in feature importance and how it affects customer churn. This could be due to:

1. An area experiencing unstable network connection issues due to poor coverage
2. A larger population in certain areas that could cause system overload on existing infrastructure.
3. Demographic issues in a certain area which are being addressed by a competing telecom company.

Thank You.
Any questions?

Author: Samuel Kiio Kyalo
GitHub: https://github.com/Samuel-Kiio/Customer-Churn-in-Telecommunication-Company