SYTRIATEL CUSTOMER CHURN PREDICTION

Billy Mwangi

Business Overview

SyriaTel is a leading provider of telecommunications services in the United States. It offers a wide range of services, including wireless, wireline, and internet services. The company has been in business for over 50 years and has a strong customer base.

Business Problem

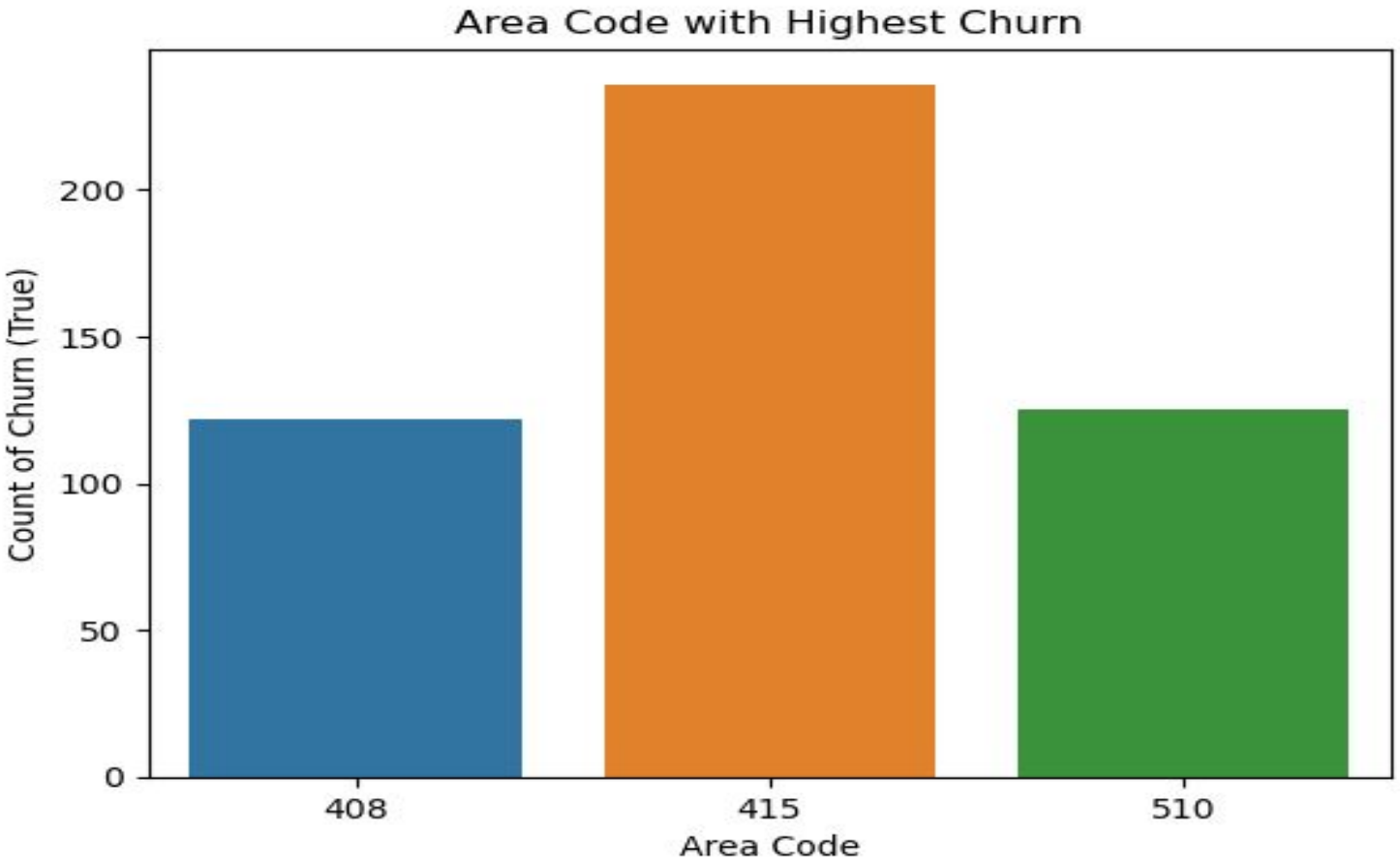
In recent years, the company has been facing an issue with customer churn. Churn is the rate at which customers cancel their service with a company. The company's churn rate has been increasing over the past few years. This is a major concern for the company because it is losing revenue and customers. The company would like to predict customers who are likely to churn. This will help the company to identify customers who are at risk of leaving and take steps to prevent them from leaving.

Objectives

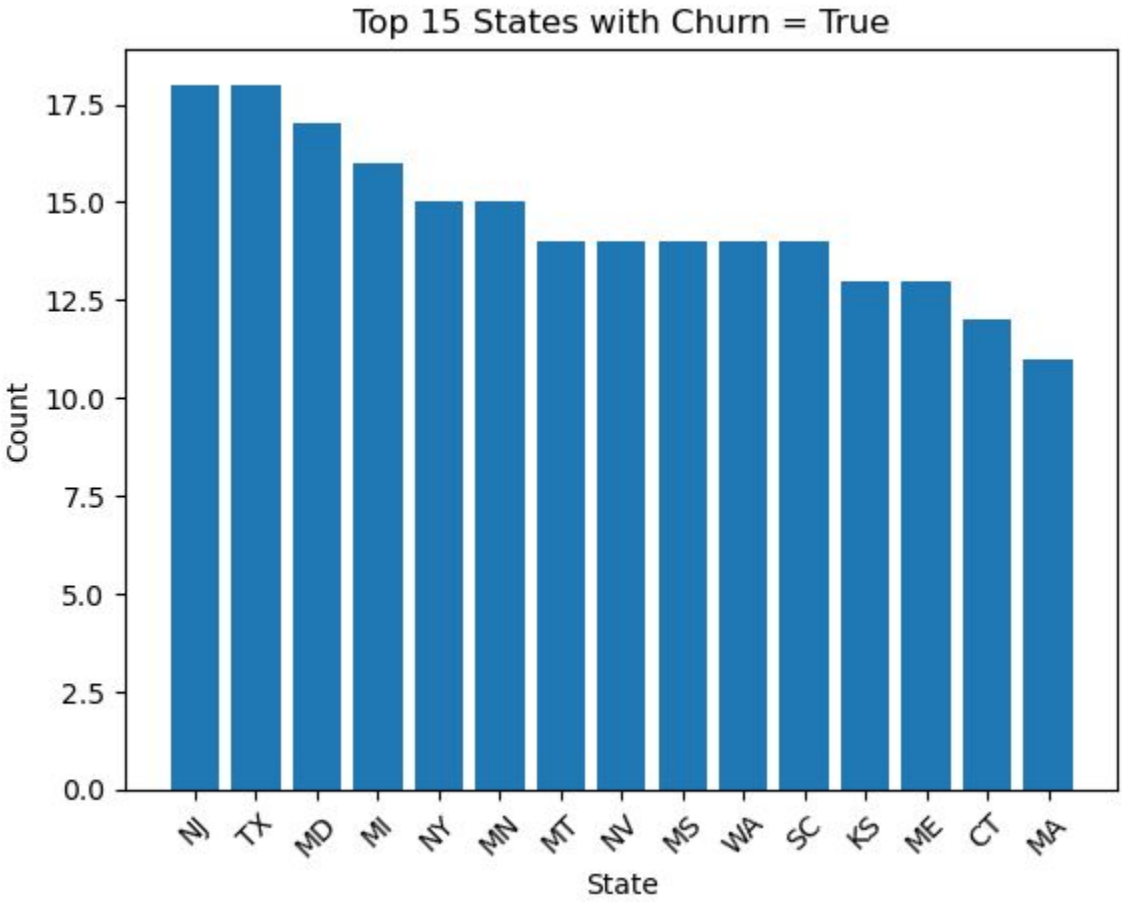
The goal of this project is to:-

- Predict customers who are likely to churn. This will help the company to identify customers who are at risk of leaving and take steps to prevent them from leaving.
- Check for relationship between various variables and churn.
- Find out the features that most predict customer churn.

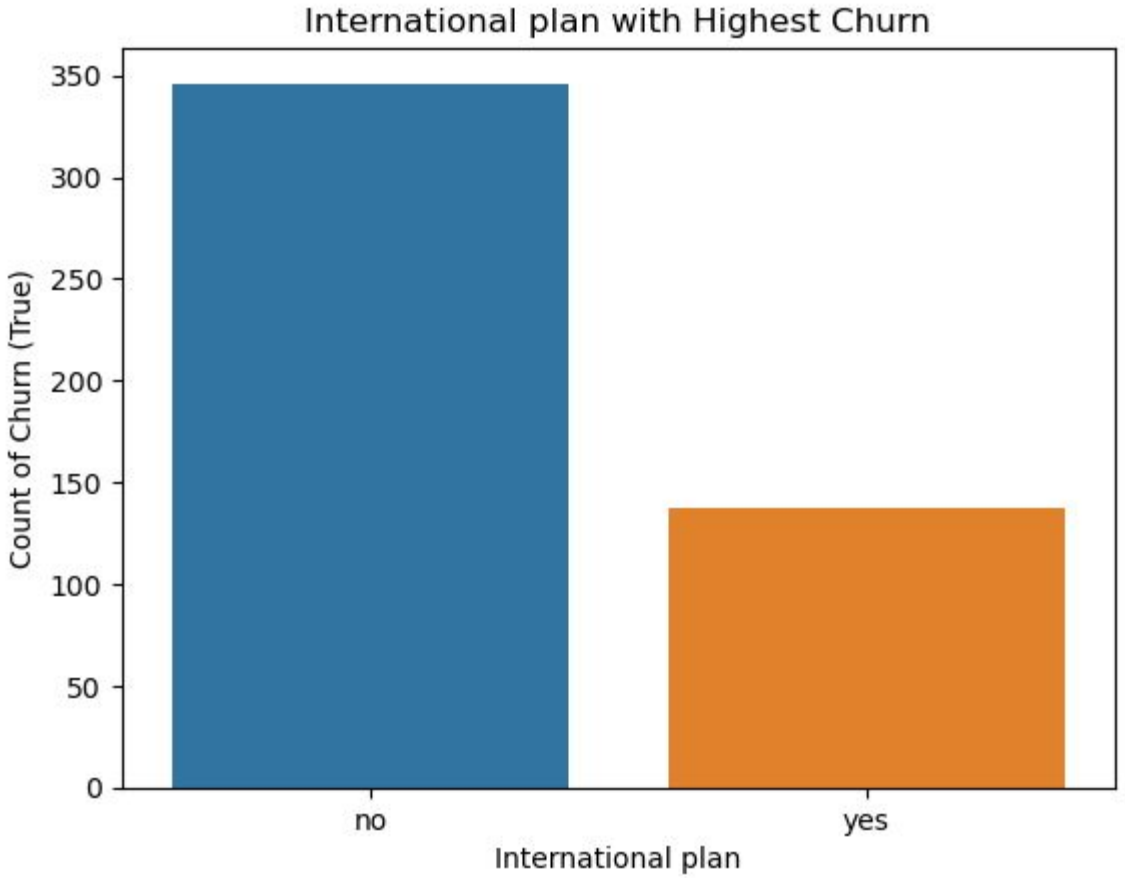
Area code with highest churn



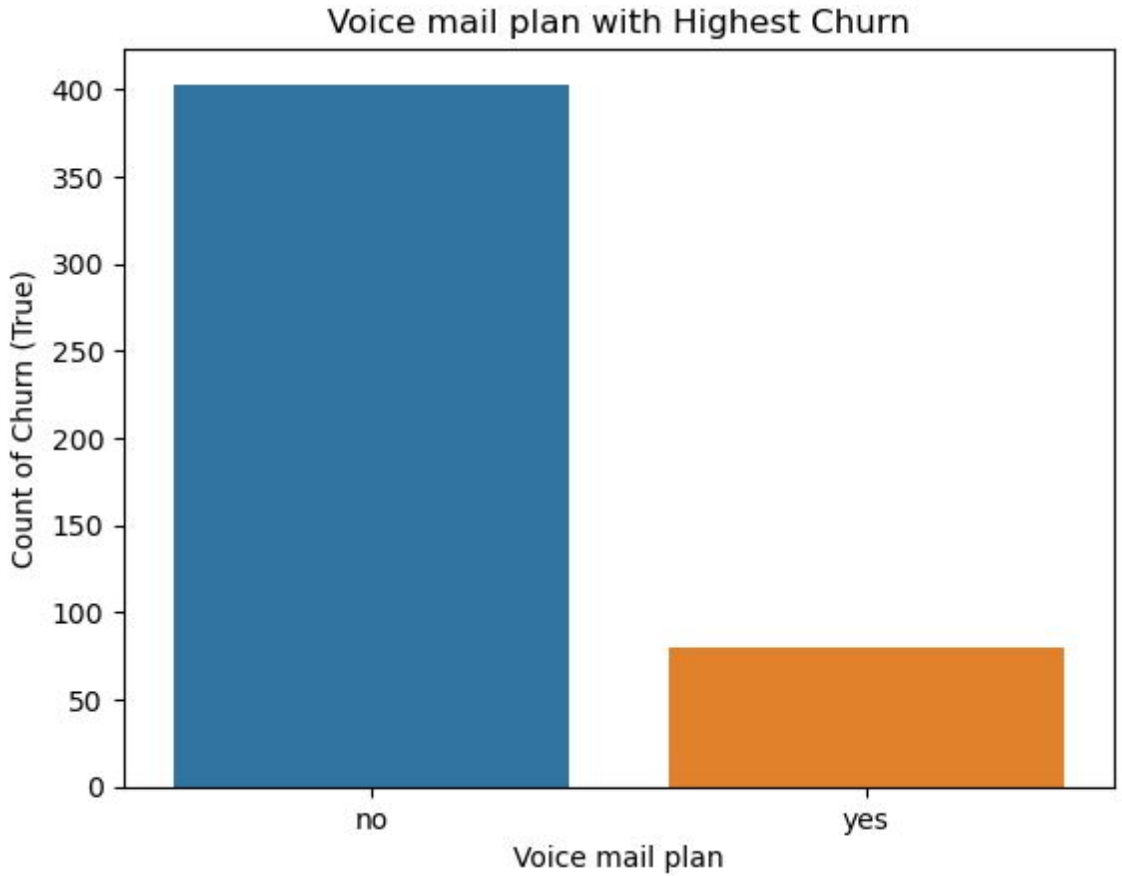
States with the highest churn



Customers with no international plan have the highest churn



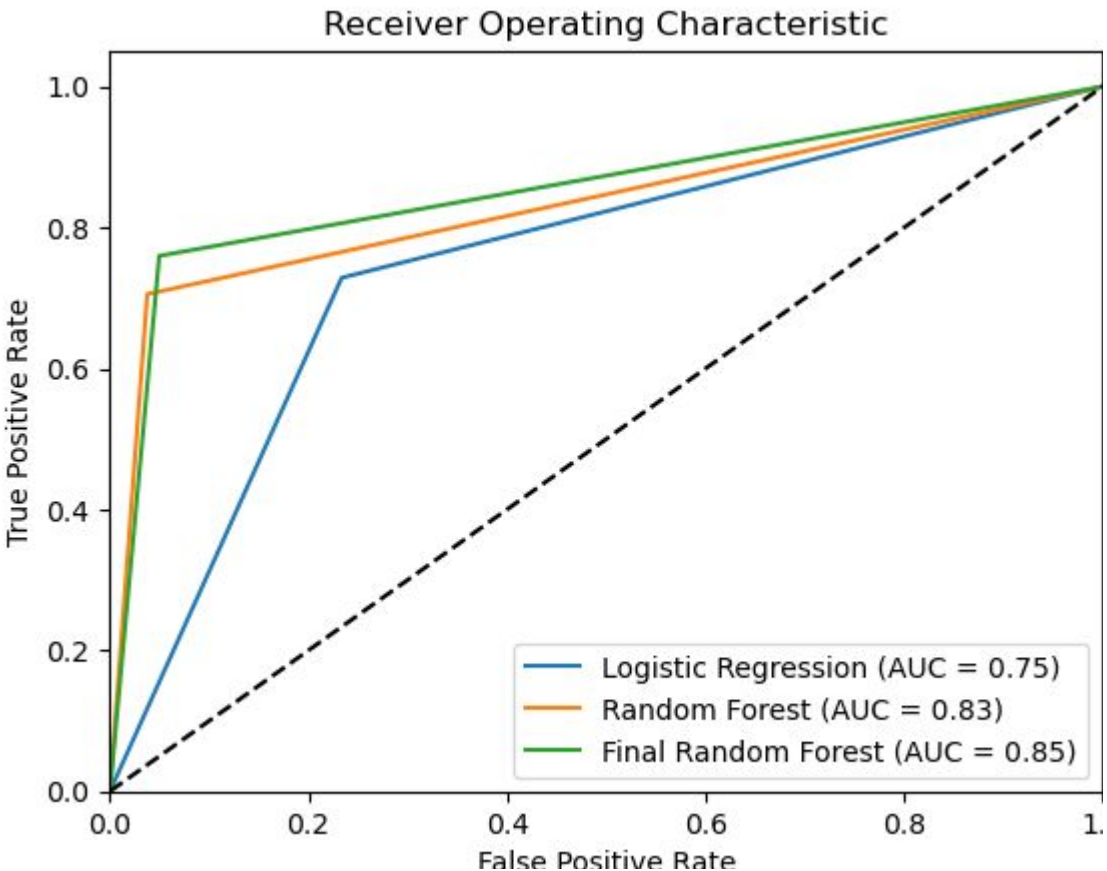
Customers with no voice mail plan have the highest churn



Modelling

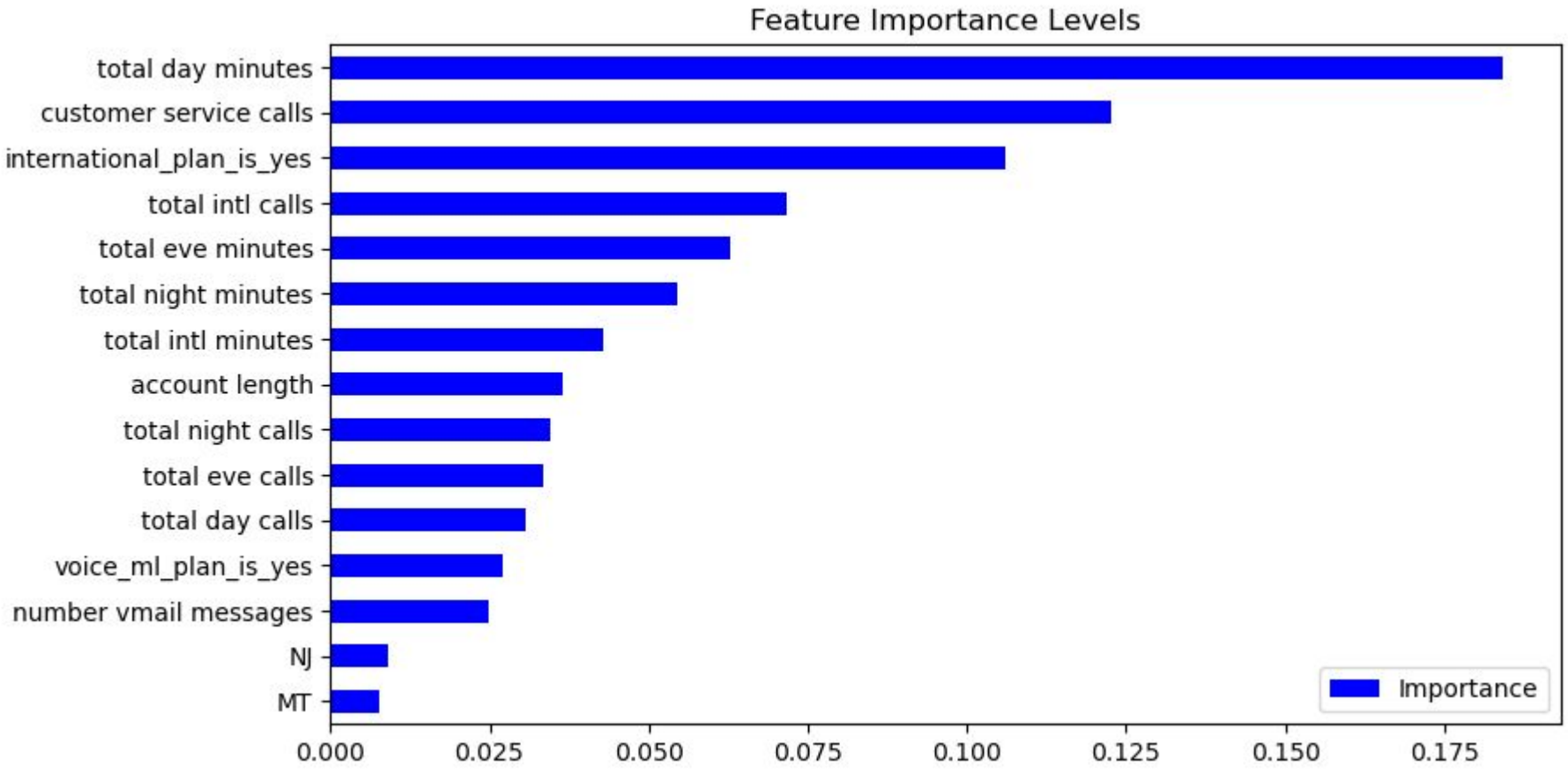
- The baseline model- logistic regression had an accuracy of 0.76 which indicates that the model is able to predict the correct outcome in about 76% of the cases, on average. The model had an F1 score of 0.497. The F1 score indicates that the model has an average level of precision and recall when predicting positive cases. The second model a random forest model had an accuracy of 0.921 and F1 score of 0.743
- The third model was a hyperparameter tuned random forest model which had a slightly decreased accuracy of 0.919 however it had a slightly improved F1 score of 0.754 indicating that the model achieves a relatively good balance between precision and recall when predicting positive cases. This suggests that the model's positive predictions are accurate, and it can effectively capture a significant portion of the actual positive cases.

Performance of the three models



- Considering the F1 score and AUC, it appears that the hyperparameter tuned random forest model (Final Random Forest model) performs better than the first model.
- AUC measures the model's ability to discriminate between positive and negative samples. The hyperparameter tuned model has a slightly higher AUC, indicating improved discrimination performance.

Features with the most predictive customer churn



Recommendations

The company should do the following things: -

- Offer customers a discount on their monthly bill if they use less minutes during peak hours. This could help to reduce the number of total day minutes that customers use, which could in turn reduce the number of customers who churn.
- Create a customer service portal where customers can easily find answers to their questions. This could help to reduce the number of customer service calls that customers make, which could in turn reduce the number of customers who churn.
- Offer customers an international plan that is more affordable. This could help to reduce the number of total international calls that customers make, which could in turn reduce the number of customers who churn.
- Send customers a text message or email reminder when they are nearing their monthly plan limits. This could help customers to be more mindful of their data usage, which could in turn reduce the number of customers who churn.
- Offer customers a loyalty program that rewards them for staying with the company. This could help to create a sense of loyalty and make customers less likely to churn.