Customer Churn Prediction – SyriaTel Telecommunications

The Team

- Benta Irungu
- Carolyne Kambura
- Harriet Joseph
- Kepha Atika

Outline

- Business Understanding
- Main Objective
- Specific Objectives
- Data Understanding
- Exploratory Data Analysis
- Modelling
- Conclusion
- Business Recommendations
- Future Steps

Business Understanding

- SyriaTel is a telecommunication company.
- The company provides services including voice and data.
- Recently, the company has been concerned about the increased rate of customer churn that is resulting in high revenue loss.
- The company is looking to outsource a data scientist to help identify the contributing factors that are leading to customers opting out of the services.
- The goal is to use this data to identify customers that are likely to churn, and take measures to keep them from doing so.

Main Objective

- To identify the factors that contribute to customer churn and develop a classifier that predicts which customers are likely to churn.

Specific Objectives

1. Conduct a comprehensive analysis of SyriaTel's customer data to identify patterns and trends that contribute to customer churn.

2. Determine which variables have the highest impact on customer churn in SyriaTel's customer base.

3. Build and test a predictive model to accurately forecast the likelihood of customer churn.

4. Evaluate the performance of the predictive model and compare it with other alternative models.

5. Identify preventive measures that SyriaTel can take to reduce customer churn and retain more customers.

Data Understanding

The data contained the following information for each customer;

- State
- Account Length
- Area Code
- Phone Number
- Whether the customer has an international plan
- Whether the customer has a voicemail plan
- Total voicemail messages.
- Voice calls related information.
- Number of calls made to customer service
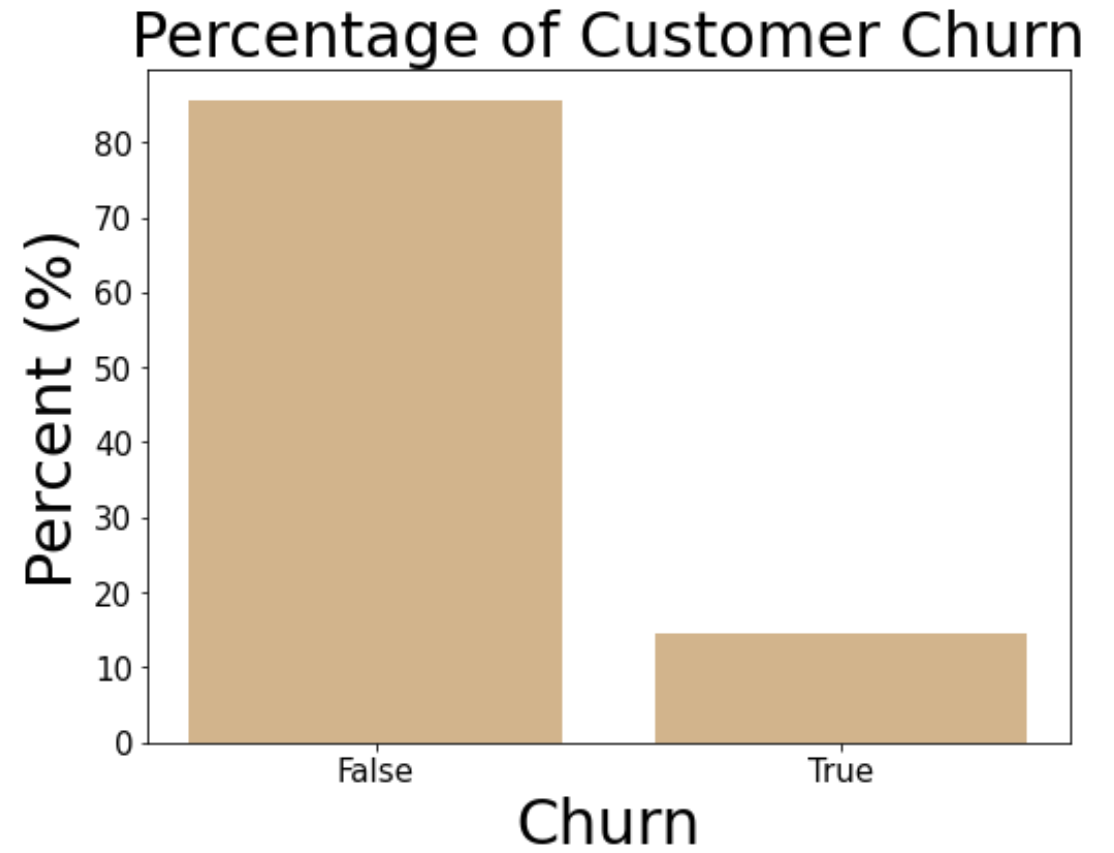- Whether a customer terminated their contract or not

Exploratory Data Analysis
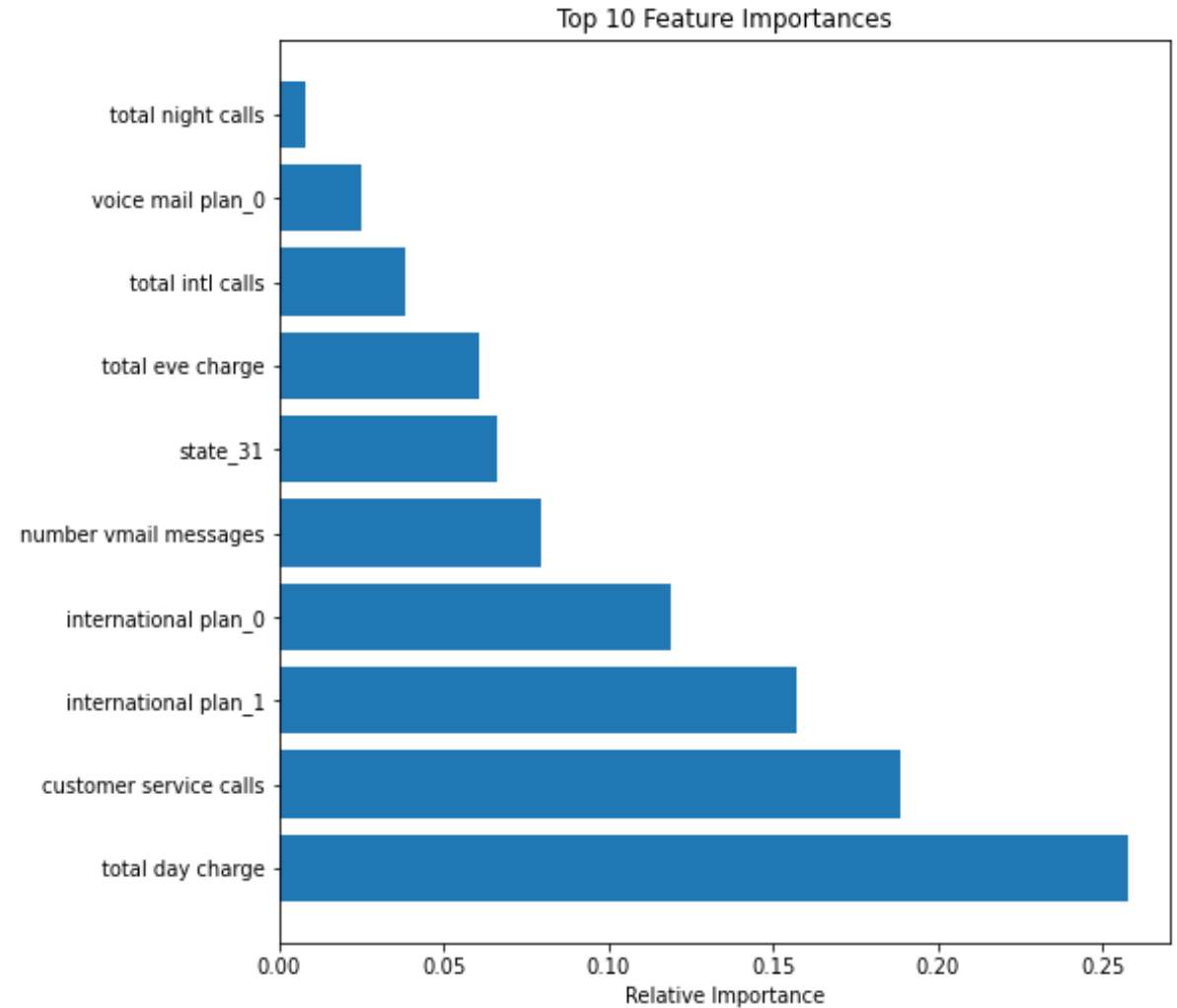
Target Variable - Churn

- As illustrated on the right, Of the 3,333 customers in the dataset, 483(14.5%) have terminated their contract with SyriaTel.

Percentage of Customer Churn

Feature Importance

Top 3 influential features for churn,

- International Plan
- Total Day Minutes
- Customer Service Calls

Top 10 Feature Importances

Analysis – Total Day Minutes

- Customers who spent more minutes on the phone during the day are more likely to churn.

- Churners spent more than 200 minutes per day on average on the phone.

Total Day Minutes for Stay vs. Churn
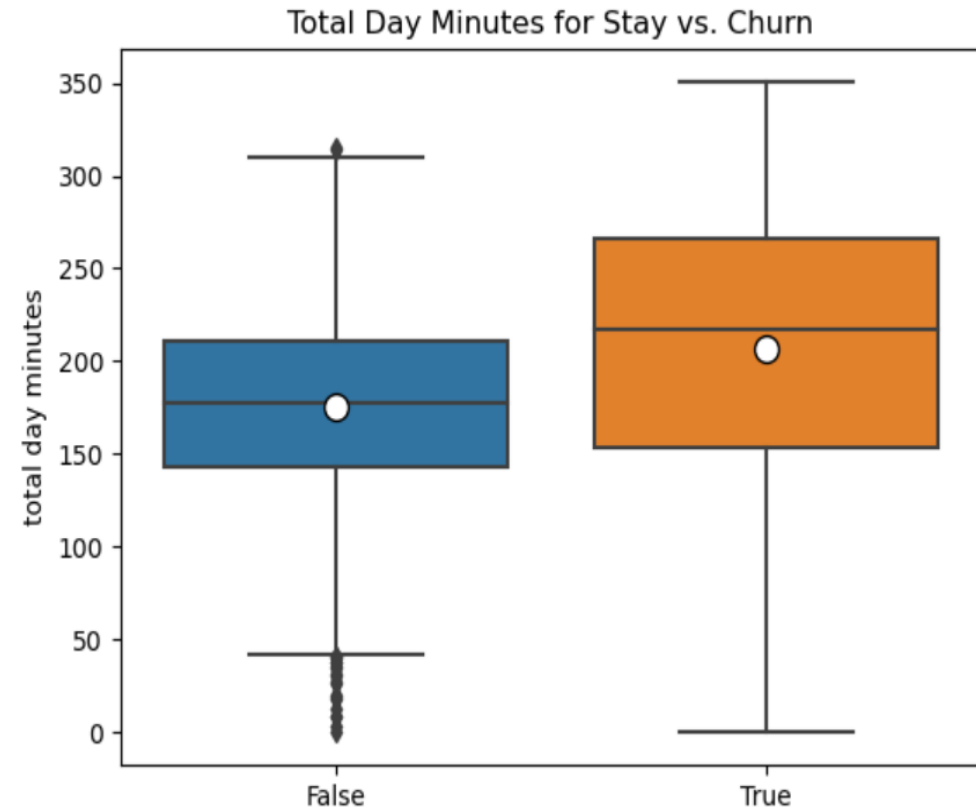
Analysis – Customer Service Calls

- Churners called customer service at least once.

Count of No. of Customer Service Calls for Churned Customers
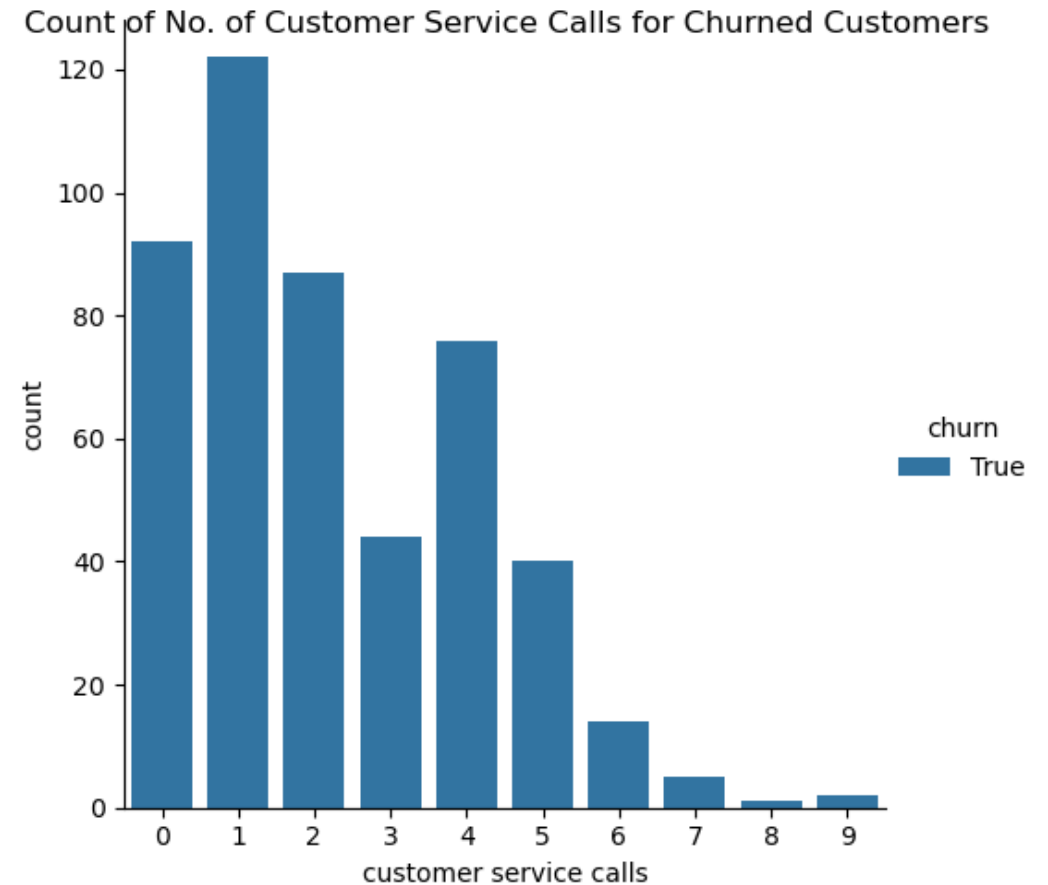
Analysis – International Plans

- The churn rate for customers with an international plan is almost four times the churn rate of customers with no international plan
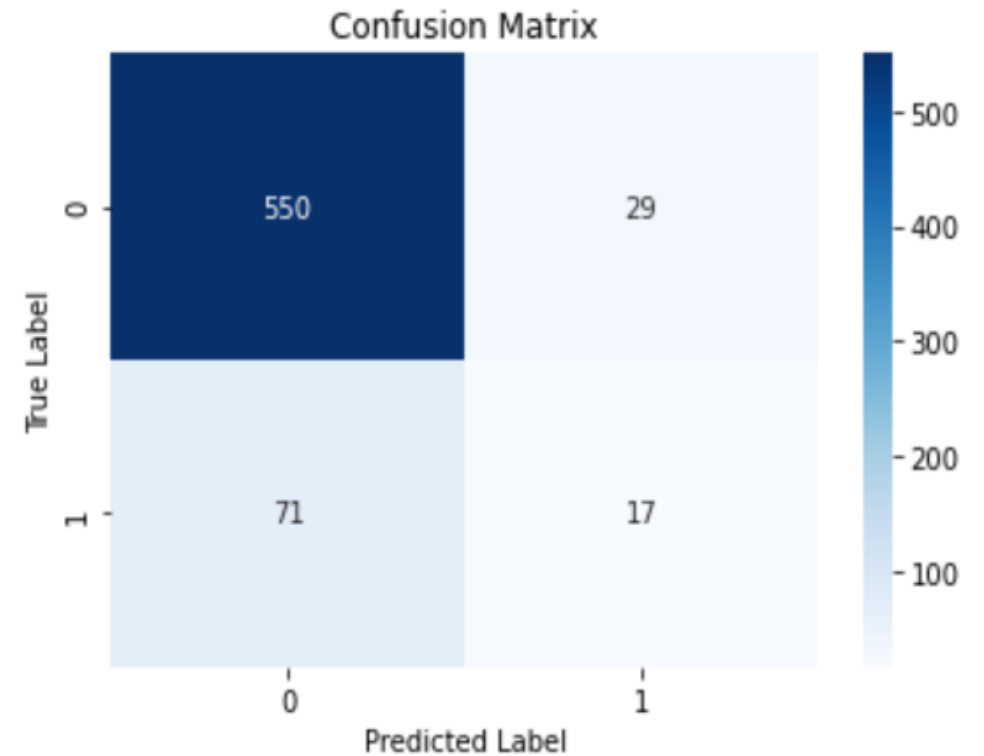

Percentage of Customer Churn for International Plan Holders
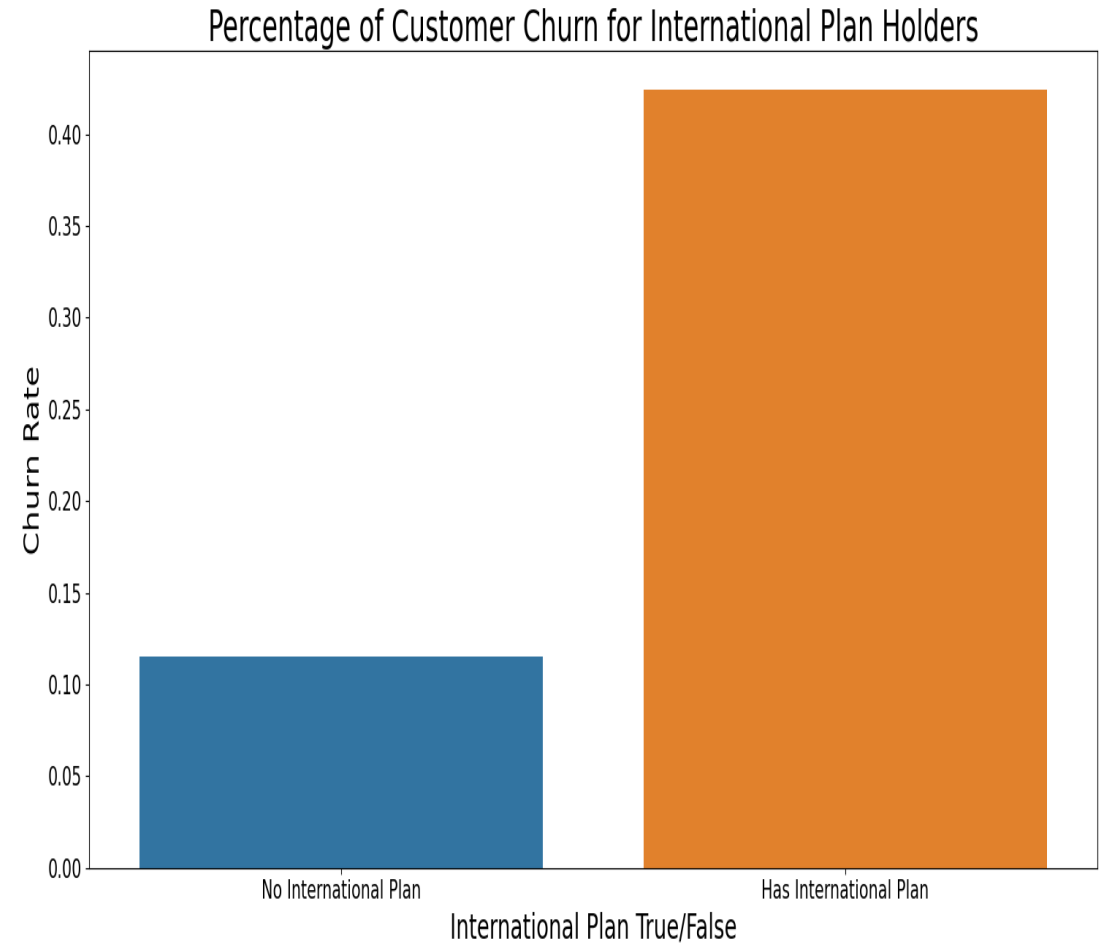
Modelling

Model 1 : K Nearest Neighbours

- False Negatives costly, therefore optimize Recall.

- Recall Score : 0.193

- 10.6% of the predictions are False Negatives which means that the model will 10.6% of the time predict that a customer will not churn yet the customer churns.


Confusion Matrix

Model 2 : Logistic Regression

- False Negatives costly, therefore optimize Recall.

- Recall Score : 0.284

- 10.4% of the predictions are False Negatives which means that the model will 10.4% of the time predict that a customer will not churn yet the customer churns.
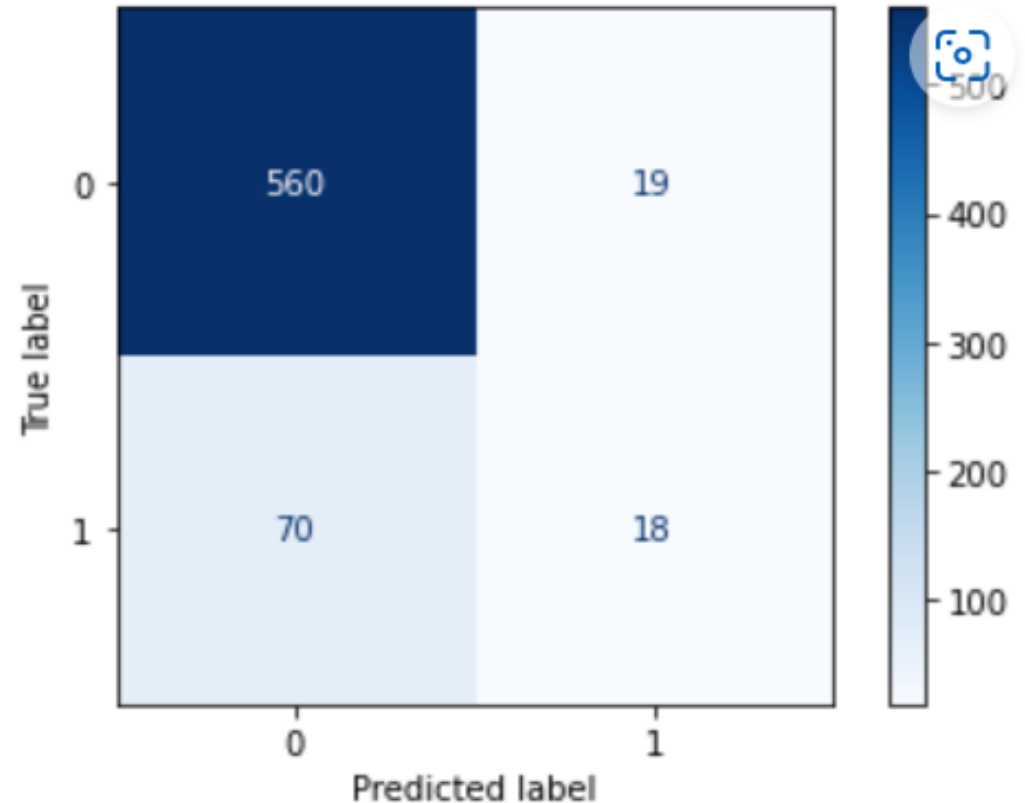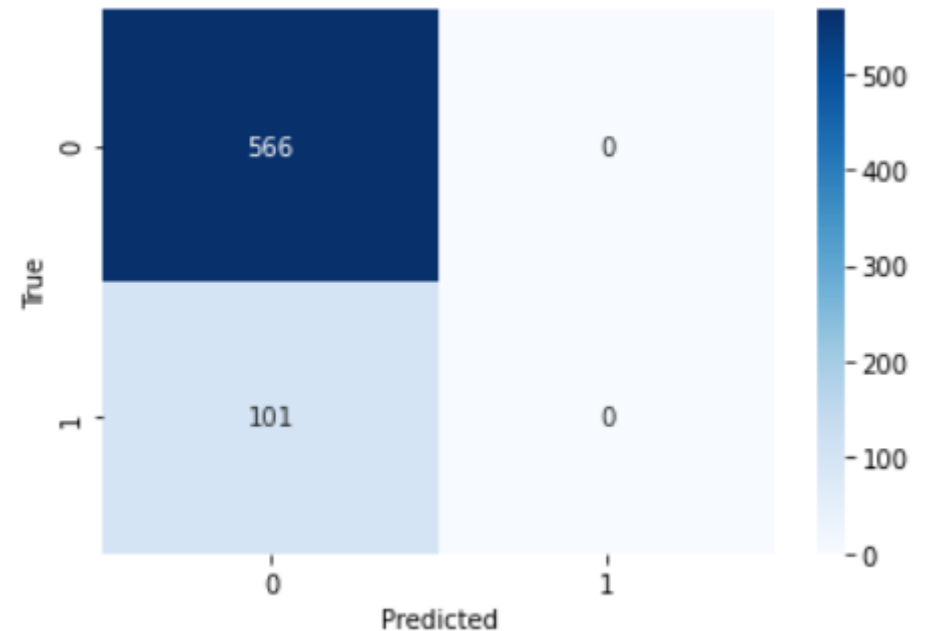
Model 3 : Decision Trees

- Overall, with a high accuracy of 0.929 and a decent mix of precision and recall, the final model with modified hyperparameters appears to be the best performing model. This model can be used to effectively anticipate which customers are likely to churn and to build suitable retention tactics, thus increasing revenue and customer happiness for SyriaTel.

Confusion matrix of the decision tree classifier
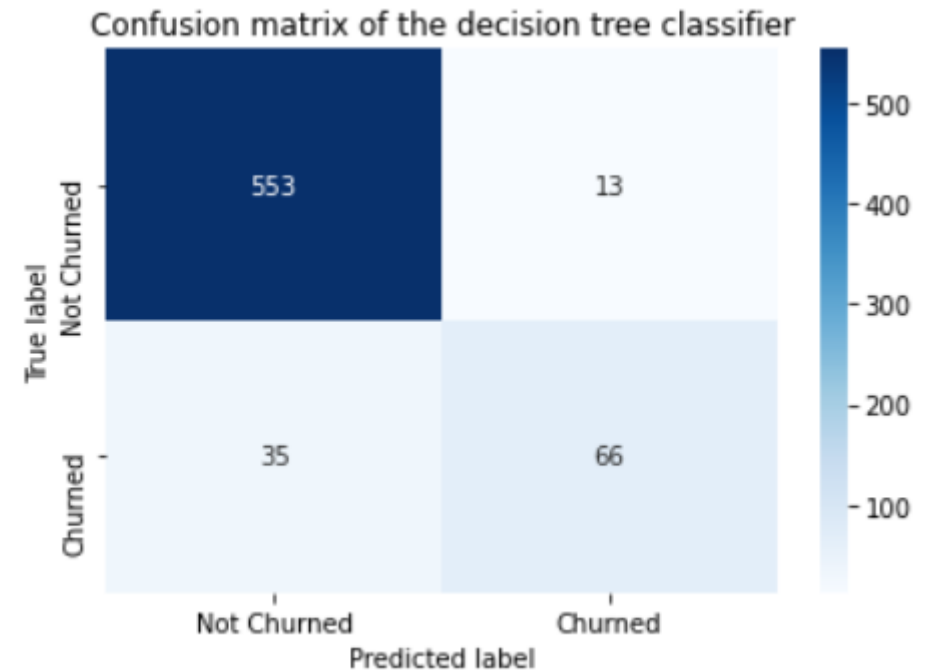
Model 4 : Random Forest

- The model accuracy, precision, recall and F1 score were 88.65%, 70%, 23.86% and 35.93% respectively.

- The model perfomed poorely than the decision tree classifier model.

Conclusion

- Importance of Recall: By prioritizing Recall, to minimize the number of customers who are incorrectly classified as non-churners.

- Best Model: the decision tree Classifier performed the best since was able to correctly identify 78% of the customers who were likely to churn. the model had an accuracy score of 92.8%

- The factors that mostly influence churn of customer include total day charge, customr service calls and number oof voice mail messages.

- Predicting customer churn is an ongoing process, and it is important to continuously refine and improve the model.

Business Recommendations

- The company should ensure continoous prediction of the factors influencing churn by contionously collecting new data and improving the model as prediction is an ongoing process.

- to reduce customer churn, the company should review the charge rate for the day calls as total day charge is the most influencial predictor for churn in this model.

- SyriaTel should improve the customer service calls through attentive listening to customers issues, feedbackd and complains and also through offering timely solutions for the same.

- SyriaTel company should reach out to the customers with high numbers of voicemail messages to determine the cause for the voicemail messages surge and know how to adress the same.

- SyriaTel should come up with a tailormade data and voice plan products for the international customers based on their unique needs.

Future Steps

1. Optimize the best model further to attain a Recall of at least 90%.

2. Explore other classification models.

3. Explore other features over and above the top 3.

4. Investigate the top 10 churn states for further insights.