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

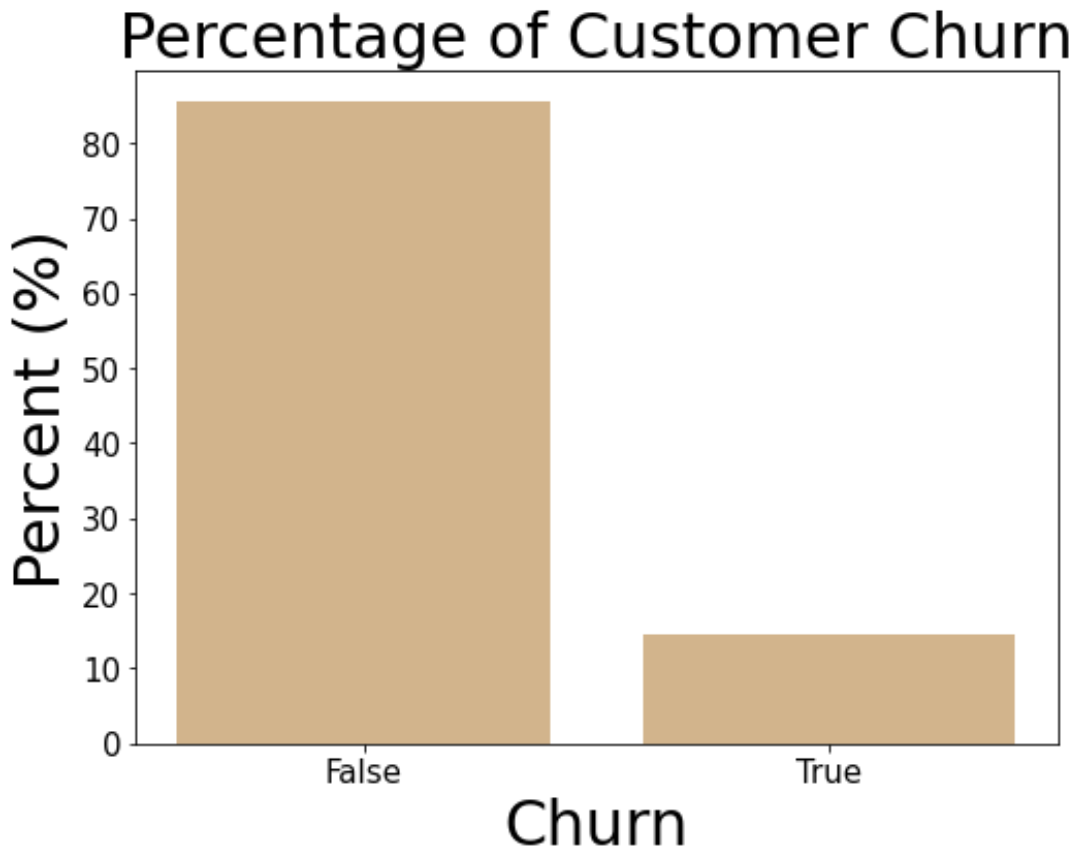
The data contained the following information for each customer;

- 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

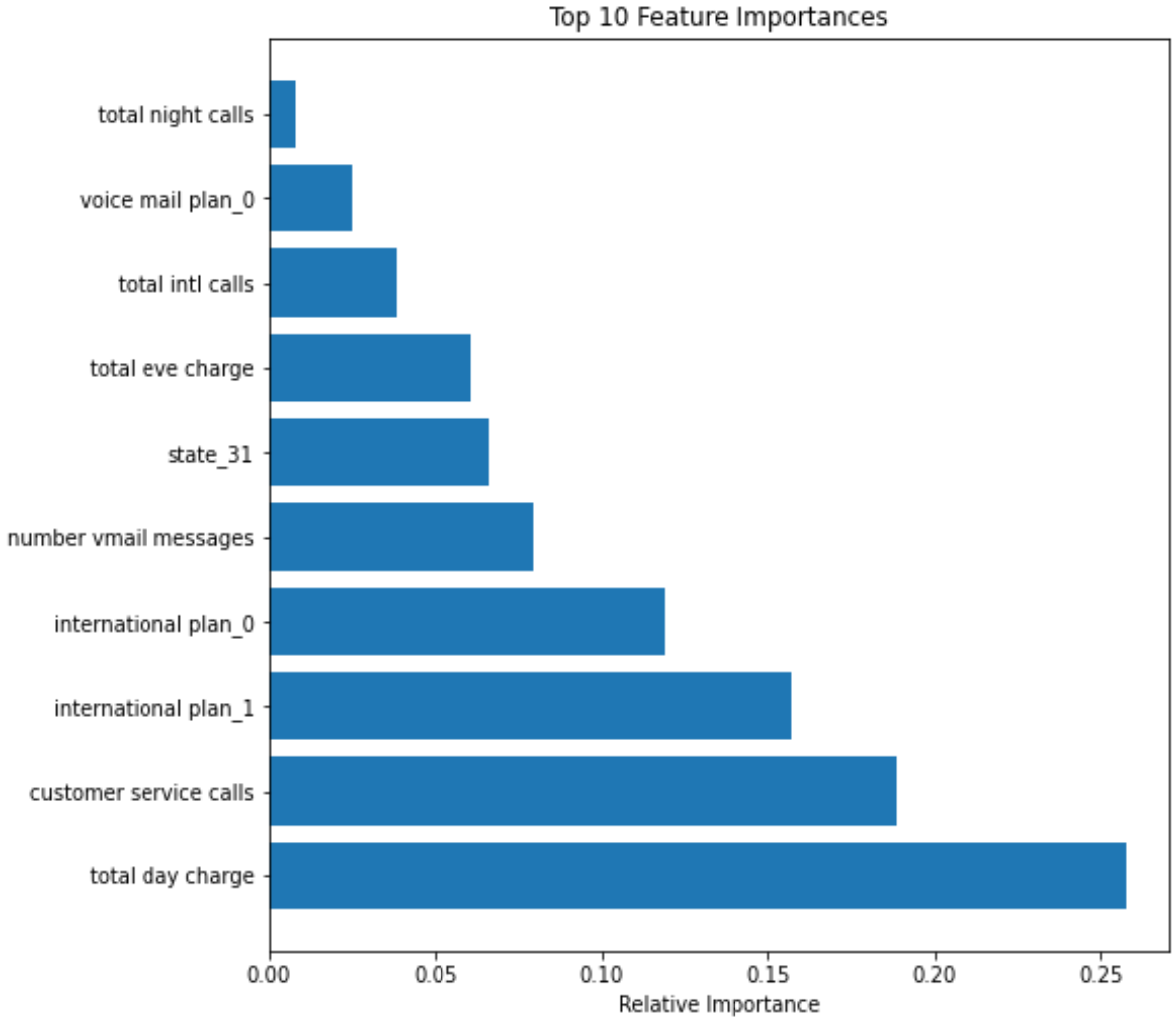
- As illustrated on the right, Of the 3,333 customers in the dataset, 483(14.5%) have terminated their contract with SyriaTel.



Feature Importance

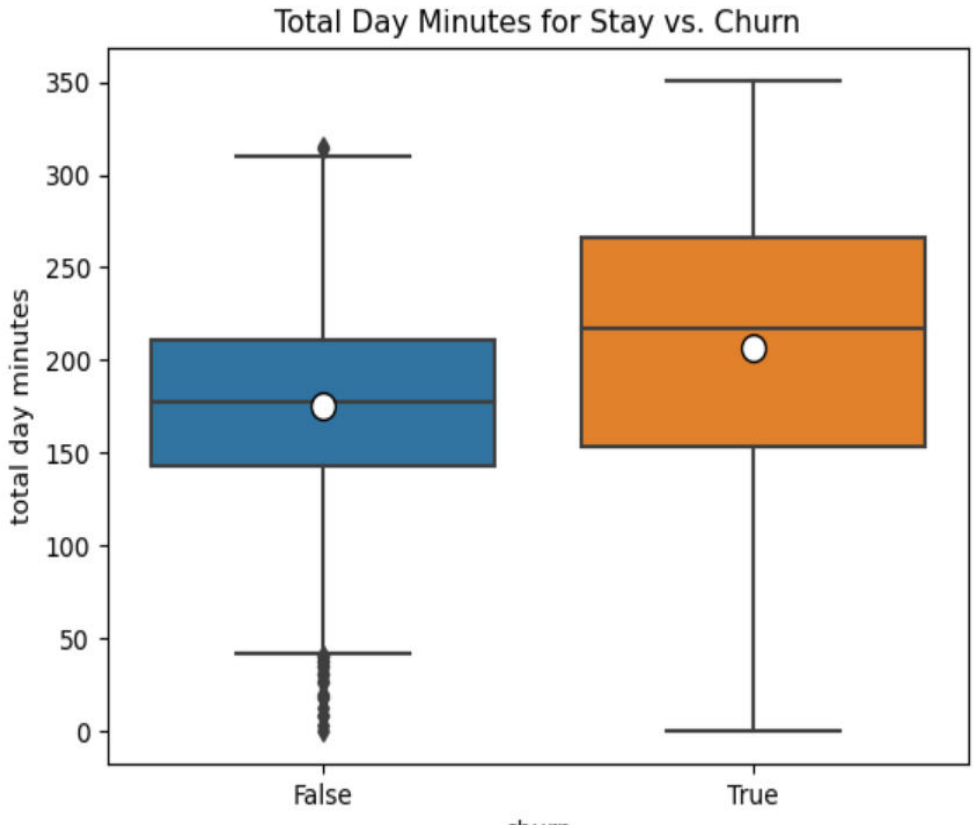
Top 3 influential features for churn,

- International Plan
- Total Day Minutes
- Customer Service Calls



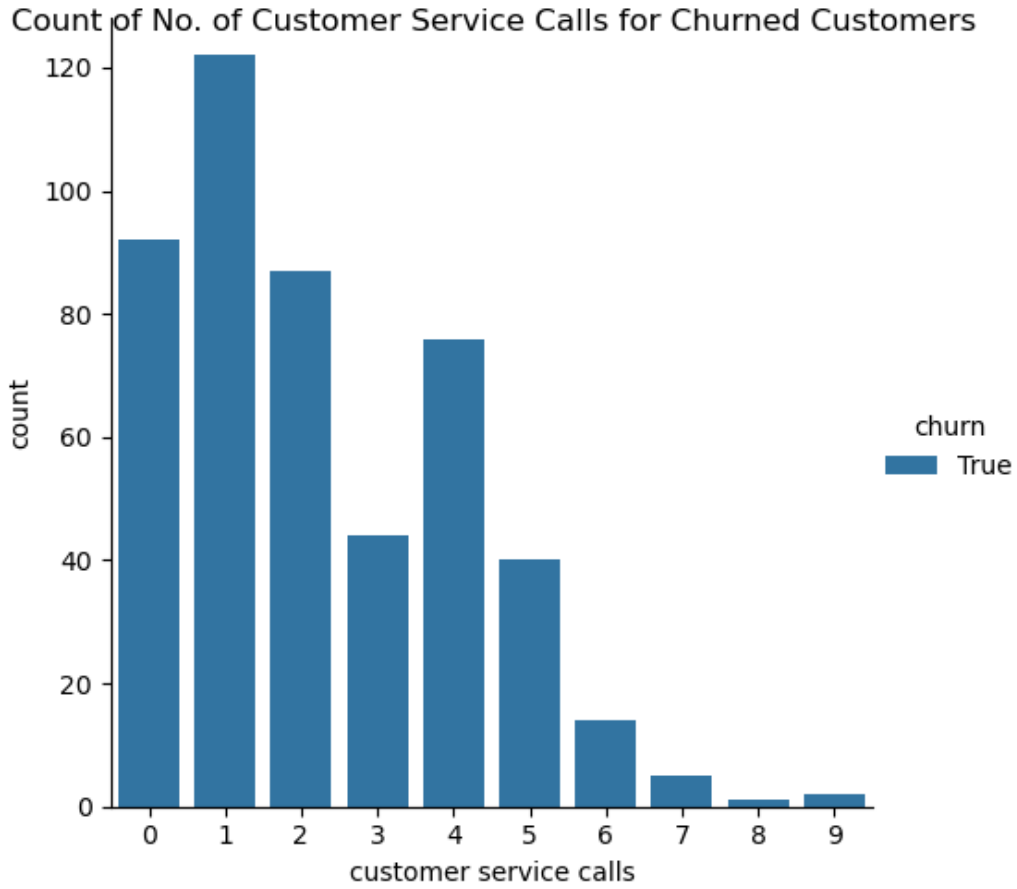
Analysis – Total Day Minutes

- 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.



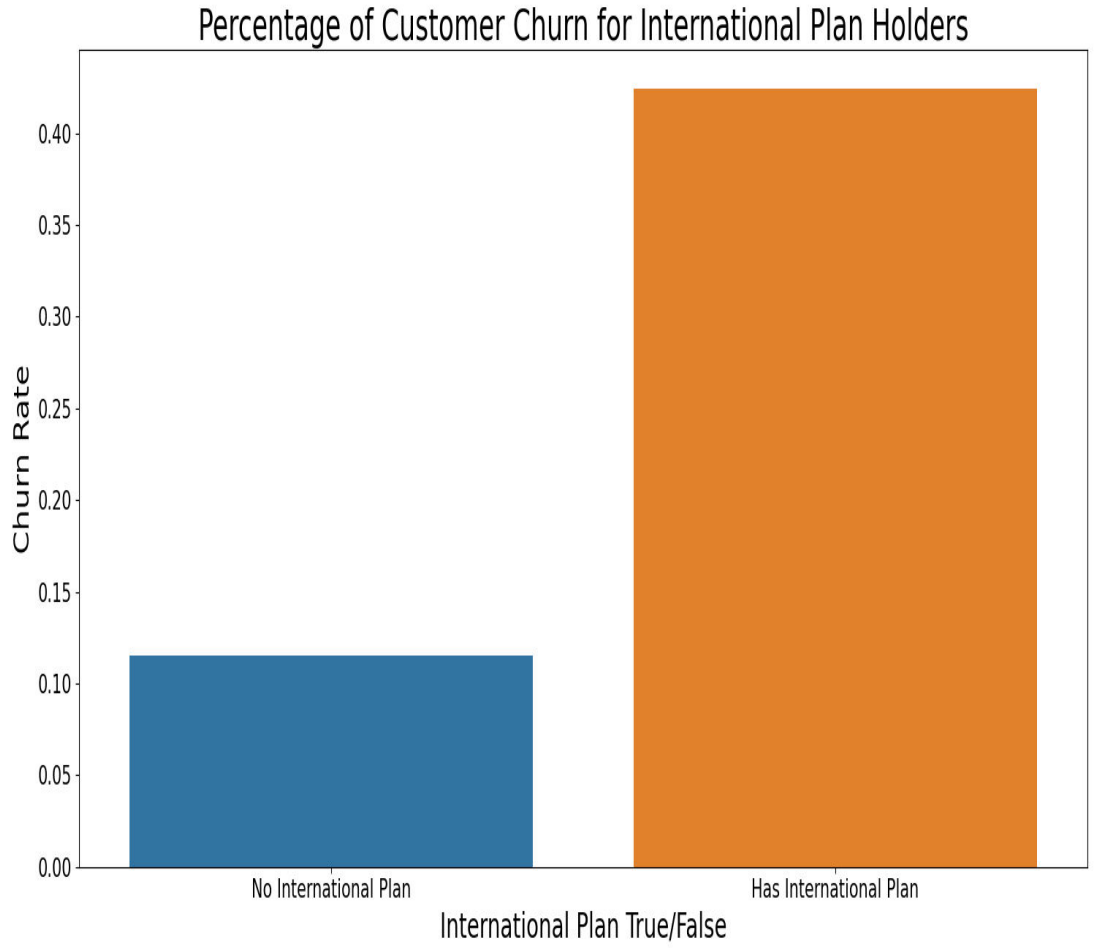
Analysis – Customer Service Calls

- Churners called customer service at least once.



Analysis – International Plans

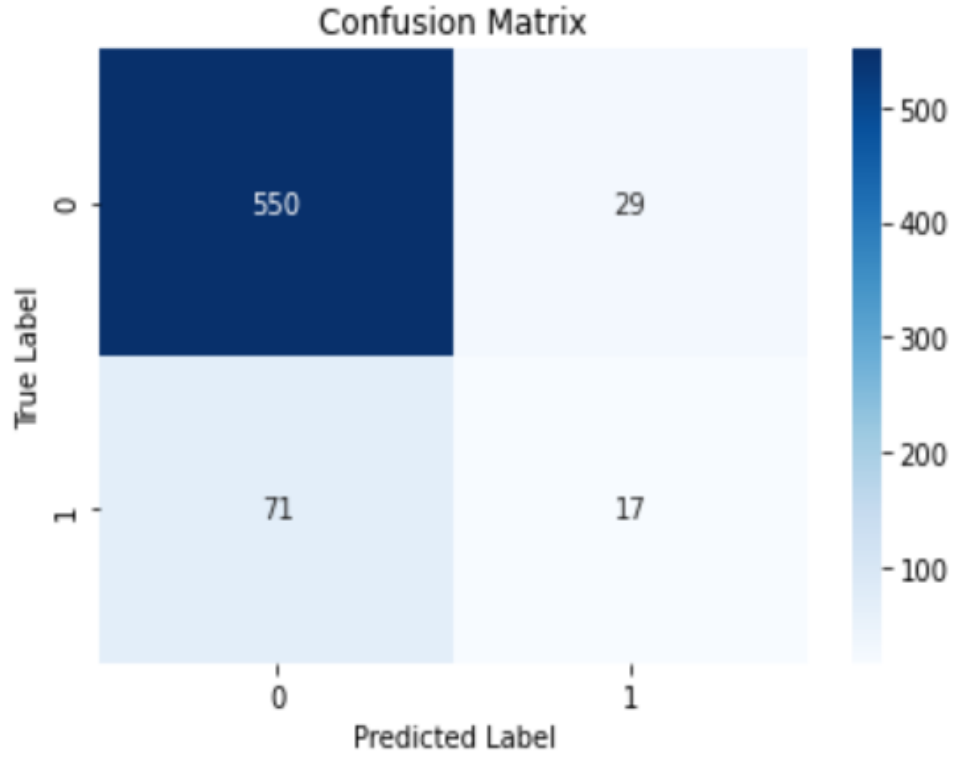
- The churn rate for customers with an international plan is almost four times the churn rate of customers with no international plan



Modelling

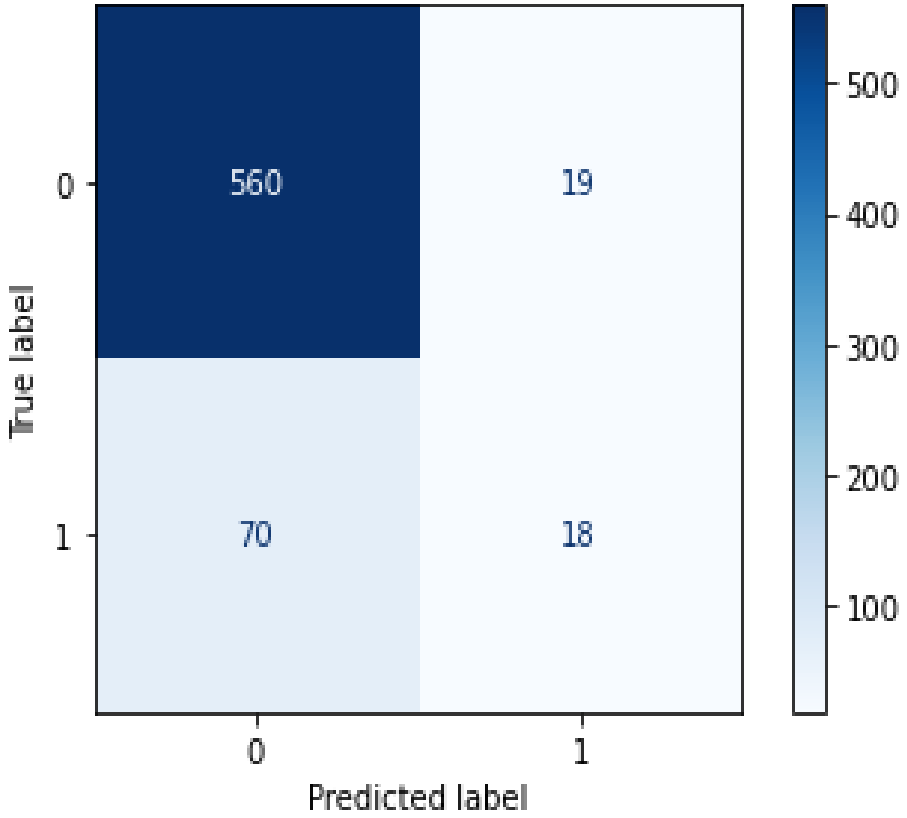
Model 1 : K Nearest Neighbours

- 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.



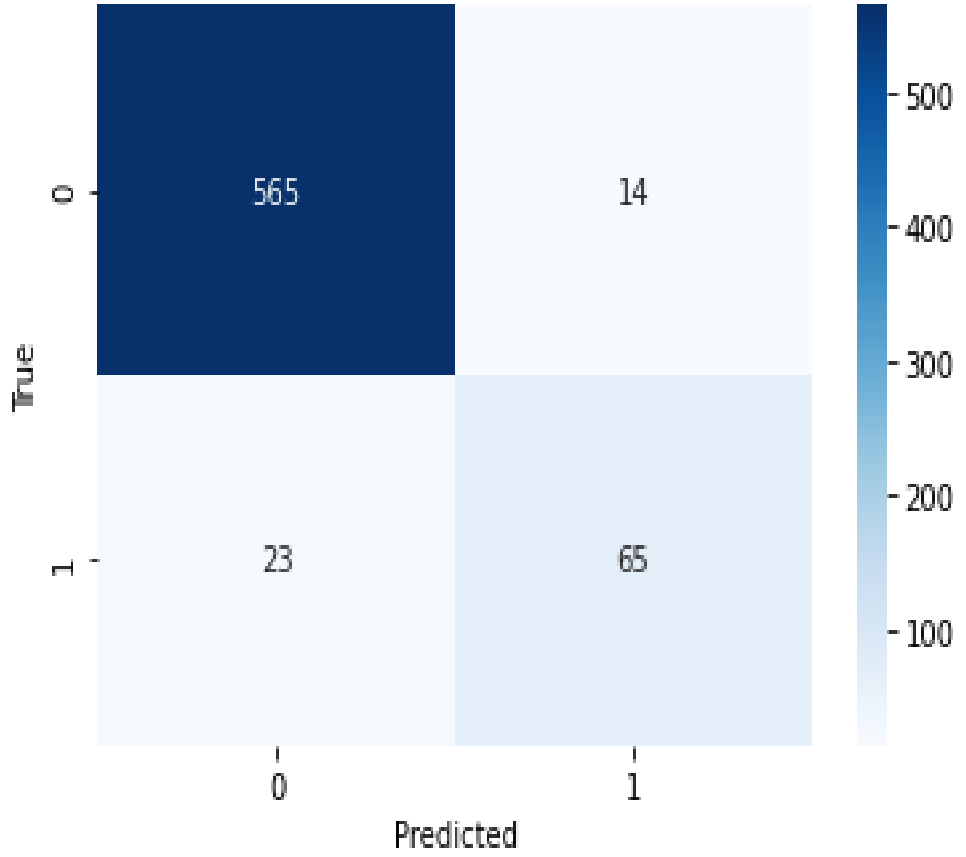
Model 2 : Logistic Regression

- 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.



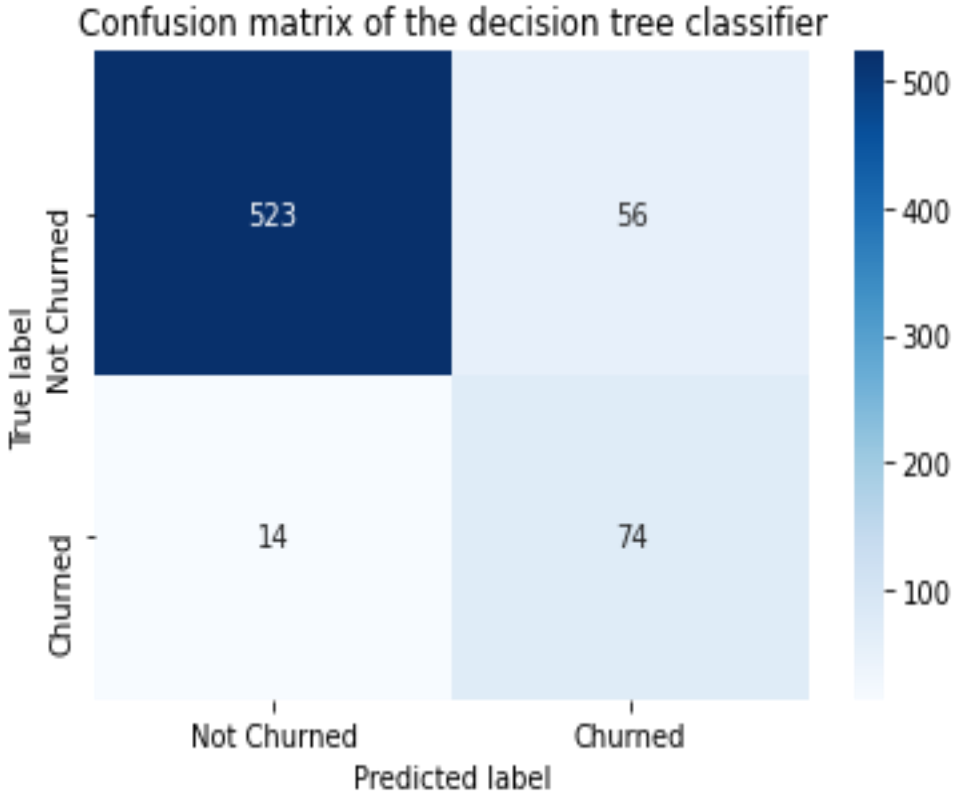
Model 3 : Random Forest

- False Negatives costly, therefore optimize Recall.
- Recall Score : 0.761
- 3.4% of the predictions are False Negatives which means that the model will 3.4% of the time predict that a customer will not churn yet the customer churns.



Model 4 : Decision Trees

- Best Model Based on Recall Metric.
- False Negatives costly, therefore optimize Recall.
- Recall Score : 0.829
- 2% of the predictions are False Negatives which means that the model will 2% of the time predict that a customer will not churn yet the customer churns.



Conclusion

- Four models were explored to determine the best model for predicting customer churn.
- False Negatives would be costly, therefore we optimized for Recall.
- The best classifier was the Decision Tree Classifier with a Recall of 0.829.
- 2% of the predictions are False Negatives which means that the model will 2% of the time predict that a customer will not churn yet the customer churns.

Business Recommendations

- Determine the unique needs of the following customers and meet them;

1. Heavy daytime callers. Come up with tariff incentives for them
2. Customers with international plans. Come up with unique retention plans
3. Customers who frequently call customer service. Assist customers proactively to reduce the need to call.

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.