Bryan Valencia

SyriaTel Churn Analysis

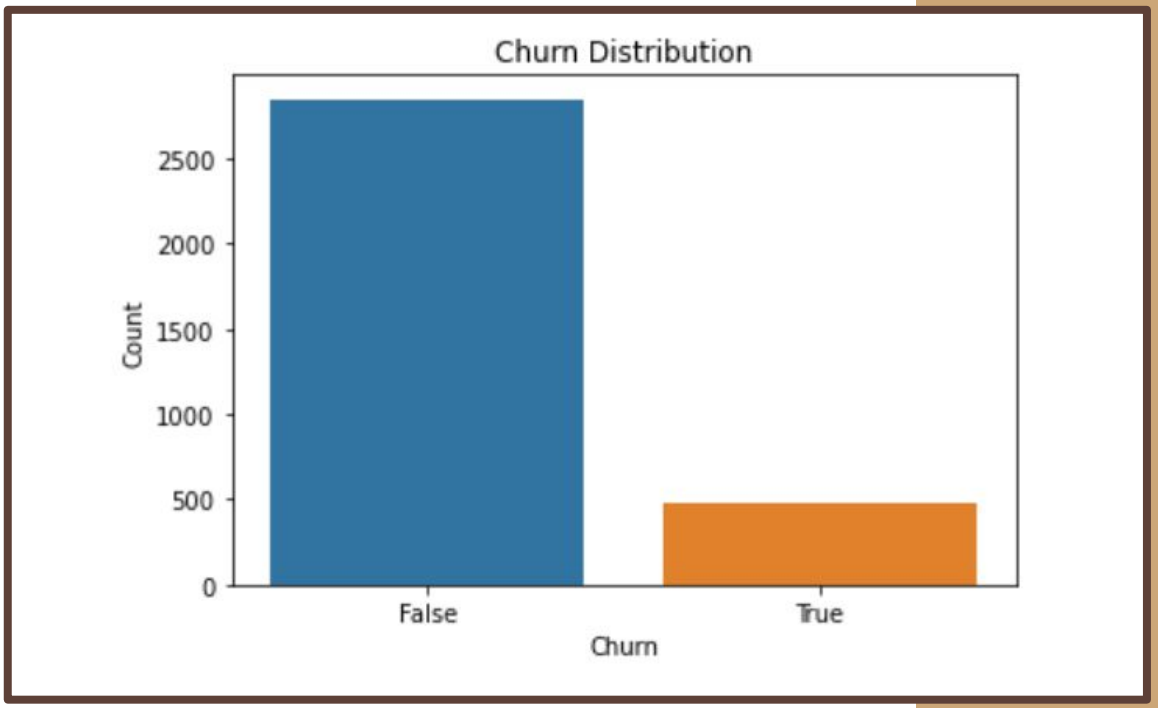
Overview

This analysis encapsules the search behind determining whether customers for a telecommunication company decide to churn or not by testing various models to see which is most accurate.

Business & Data Understanding

- SyriaTel, the stakeholder, is looking to identify what customers are likely to churn in order to take an active approach to try and retain these customers in order to prevent loss in revenue.
- By knowing what customers are likely to churn they can identify what aspects of their company need improvement.

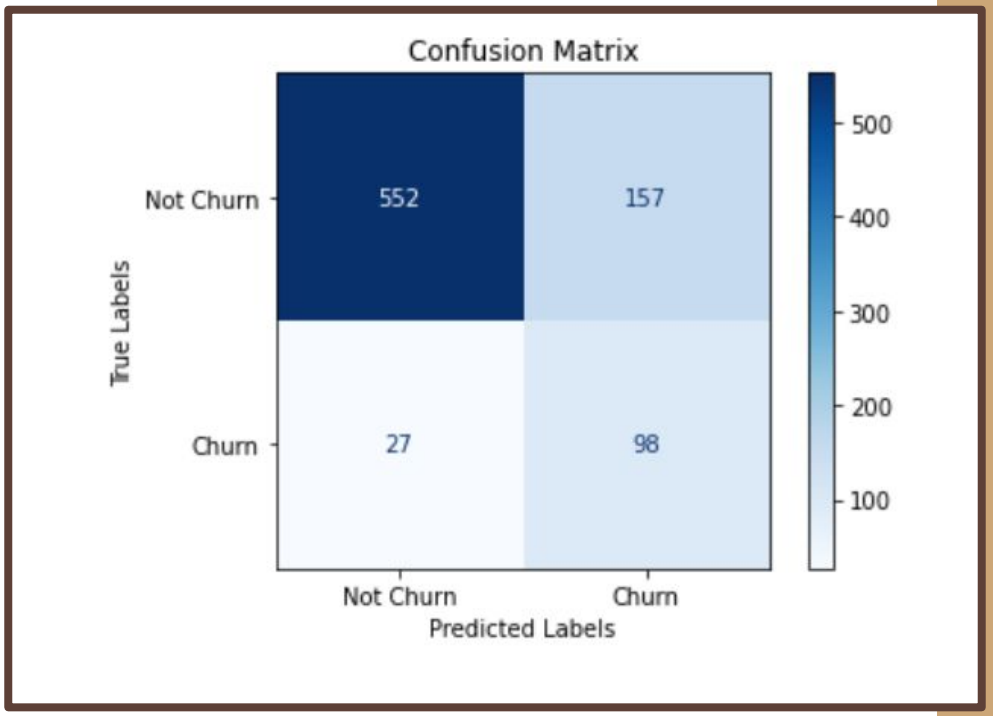
Understanding the Target Variable



- Total = 3333
- Churn = 483

- Approx. 14.5% of customers churned

Logistic Regression

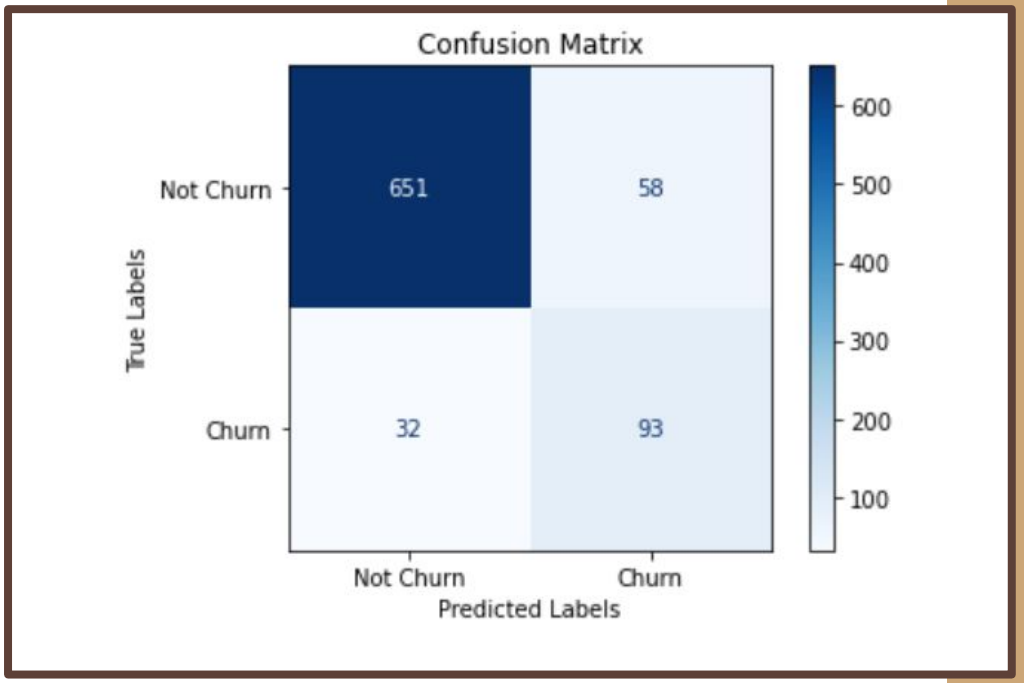


- Accuracy: 78.3%
- F1-Score: 0.51

- Excessive amount of False-Positives

- Sub-par performance

Decision Tree

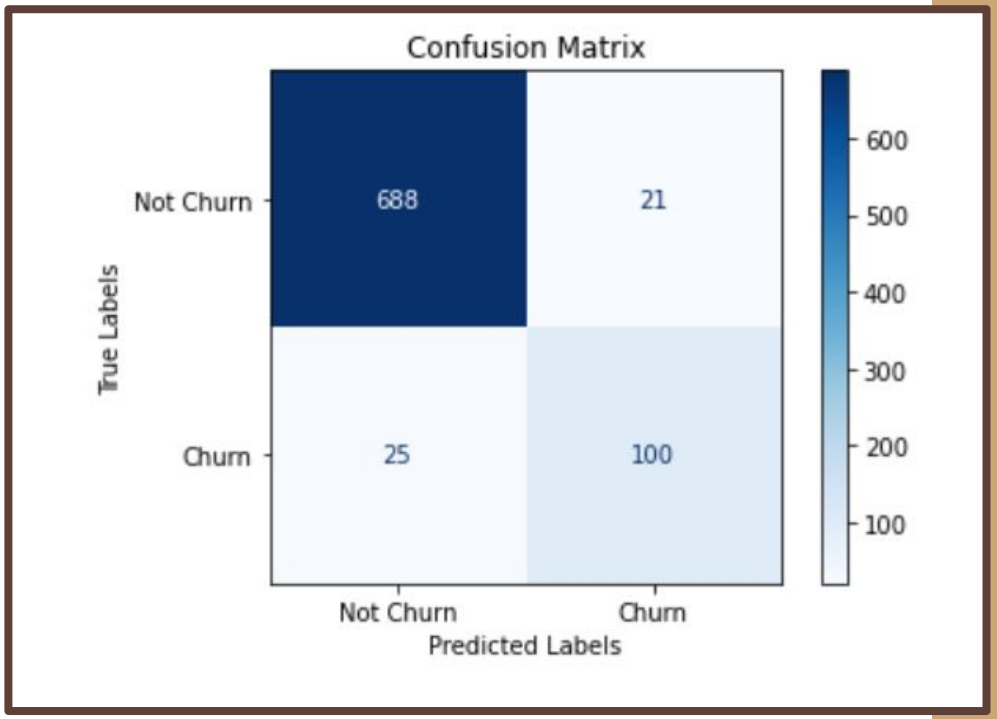


- Accuracy: 89.2%
- F1-Score: 0.67

- Contains some overfitting

- Better in comparison to Logistic Regression

Random Forest

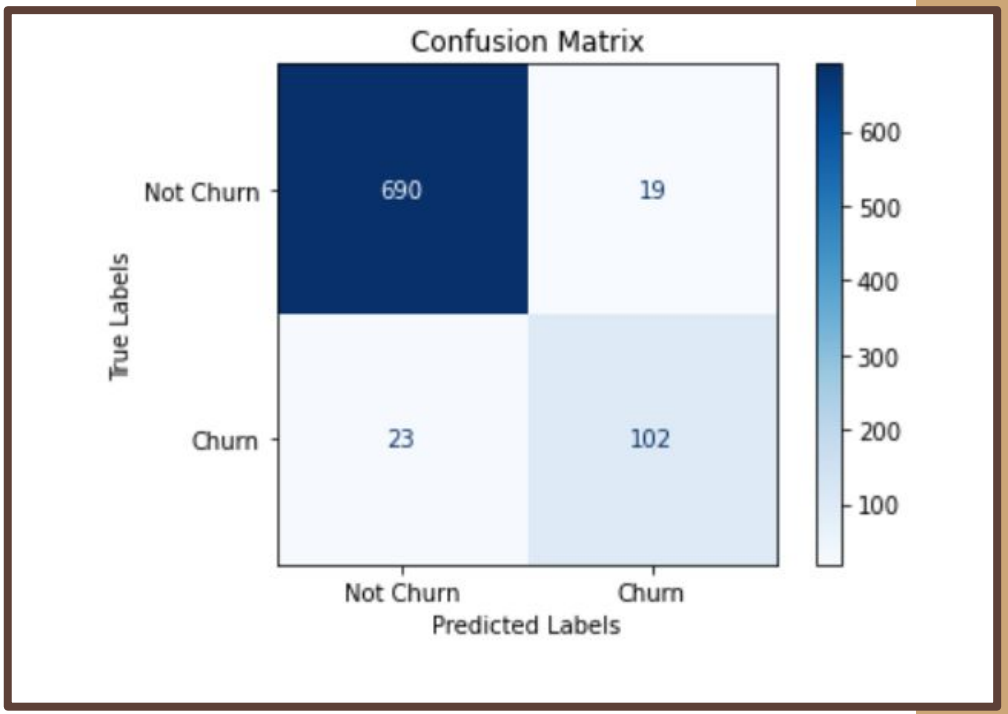


- Accuracy: 94.4%
- F1-Score: 0.81

- Still overfitting

- Improvement over Decision Tree

Random Forest w/ Hyperparameters

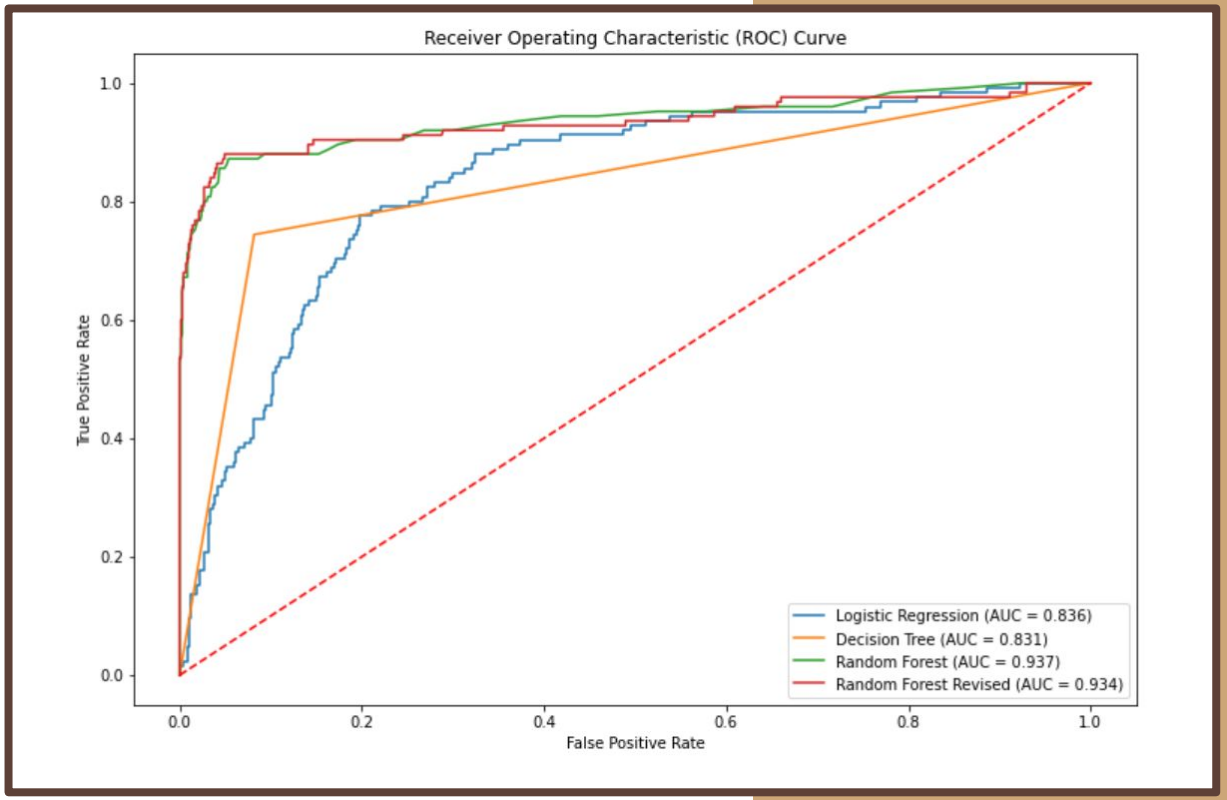


- Accuracy: 94.9%
- F1-Score: 0.82

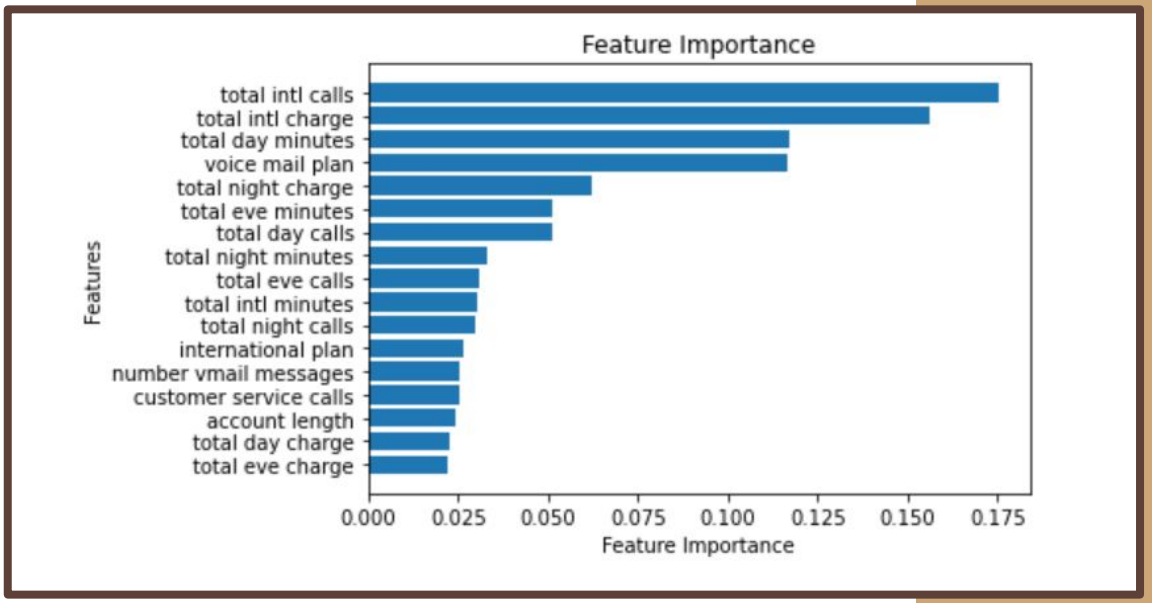
- Dealt with the overfitting

- Most successful model yet

Evaluation


Receiver Operating Characteristic (ROC) Curve
True Positive Rate
False Positive Rate
0.0
0.2
0.4
0.6
0.8
1.0
Logistic Regression (AUC = 0.836)
Decision Tree (AUC = 0.831)
Random Forest (AUC = 0.937)
Random Forest Revised (AUC = 0.934)

Evaluation



- Look into amount of international calls made and their charge

Recommendations

- Use Random Forest model with applied hyperparameters to predict churning customers
- Look into their charges and see if more affordable options can be pursued (especially internationally)

Recommendations

- Prioritize on customers that are predicted to churn first
- Review voicemail plan and company customer service

Any Questions?