SyriaTel Customer Churn

BUSINESS UNDERSTANDING

- The goal of this project is to find any predictable patterns in this data that might indicate a customer is about to churn.
- It identifies patterns and provides SyriaTel with insights that could help them improve

OBJECTIVES

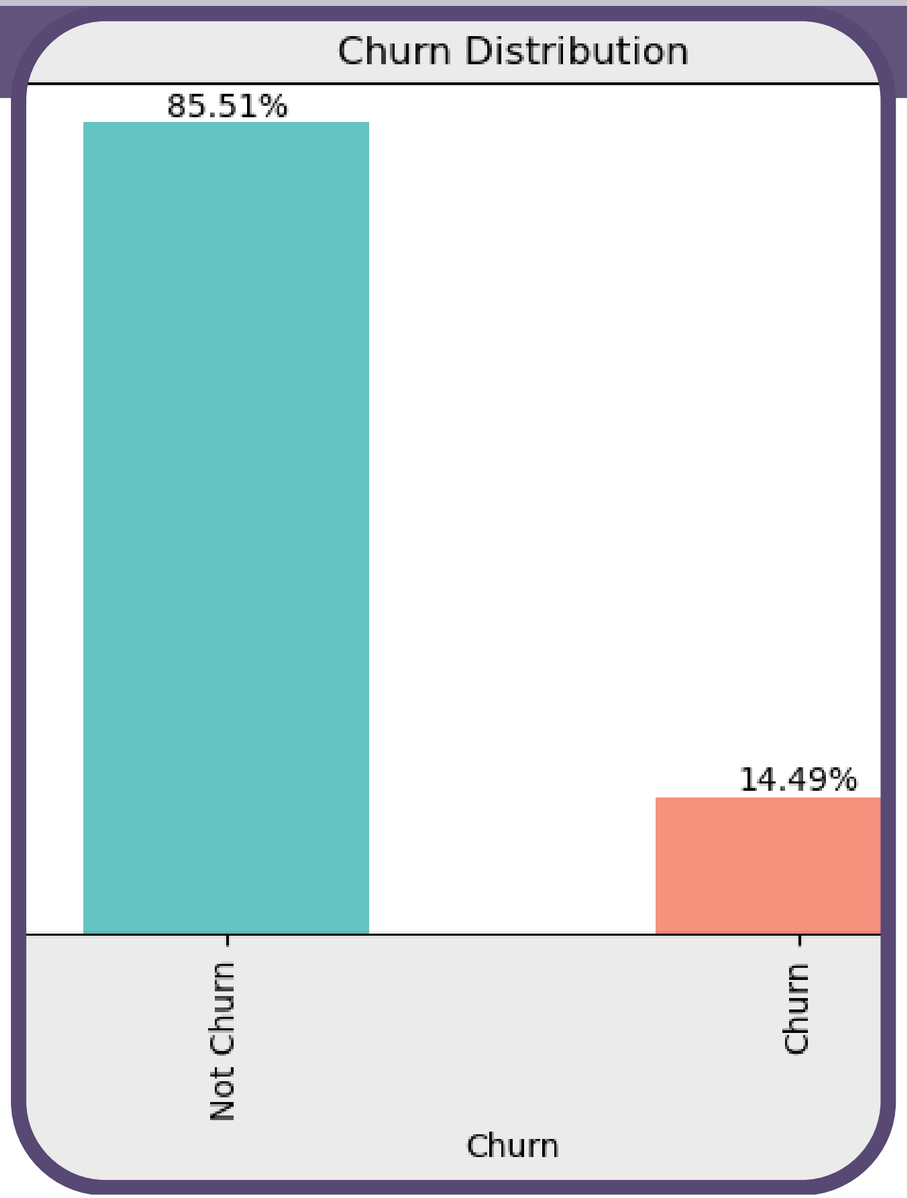
1.Identify Predictive Patterns

2.Build a Churn Prediction Model

3.Provide Actionable Insights

DATA UNDERSTANDING

- This data set is drown from Kaggle.
- The dataset contains 3333 rows and 21 columns
-



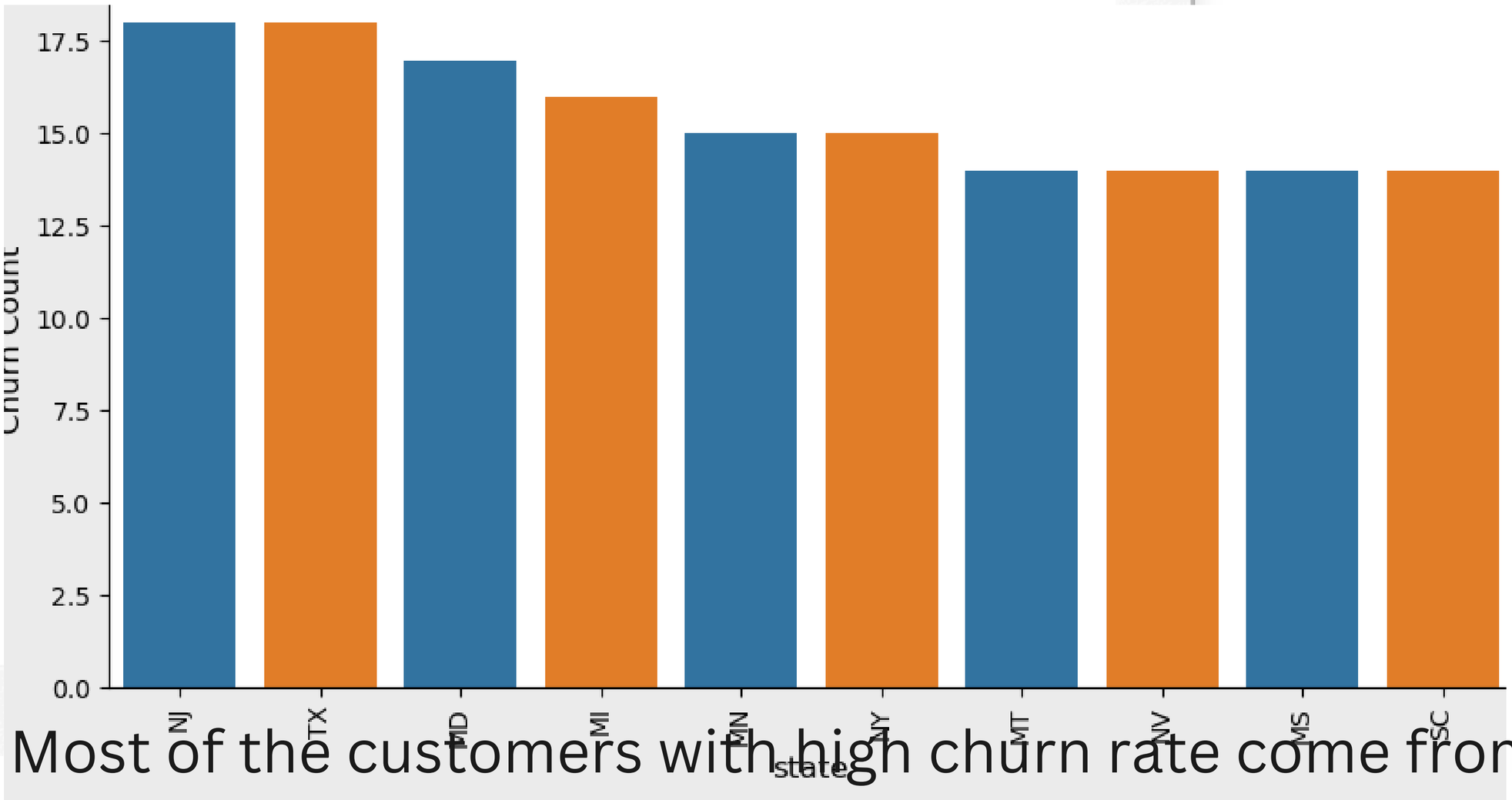
DATA ANALYSIS

Univariate Data Analysis

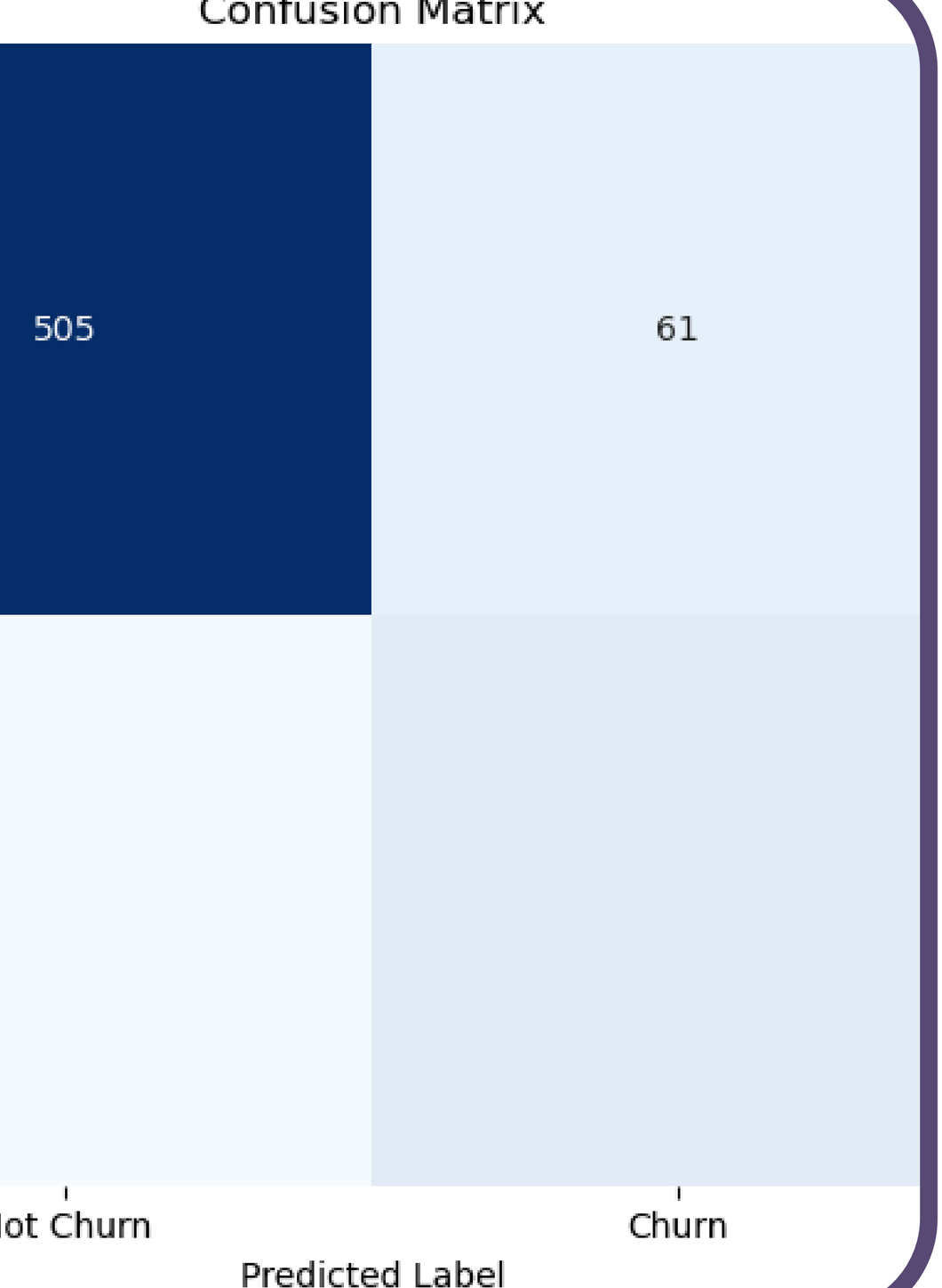
Churned customers 483

Non churned customers 2850

DATA ANALYSIS



Most of the customers with high churn rate come from New Jersy,Texas and Meryland



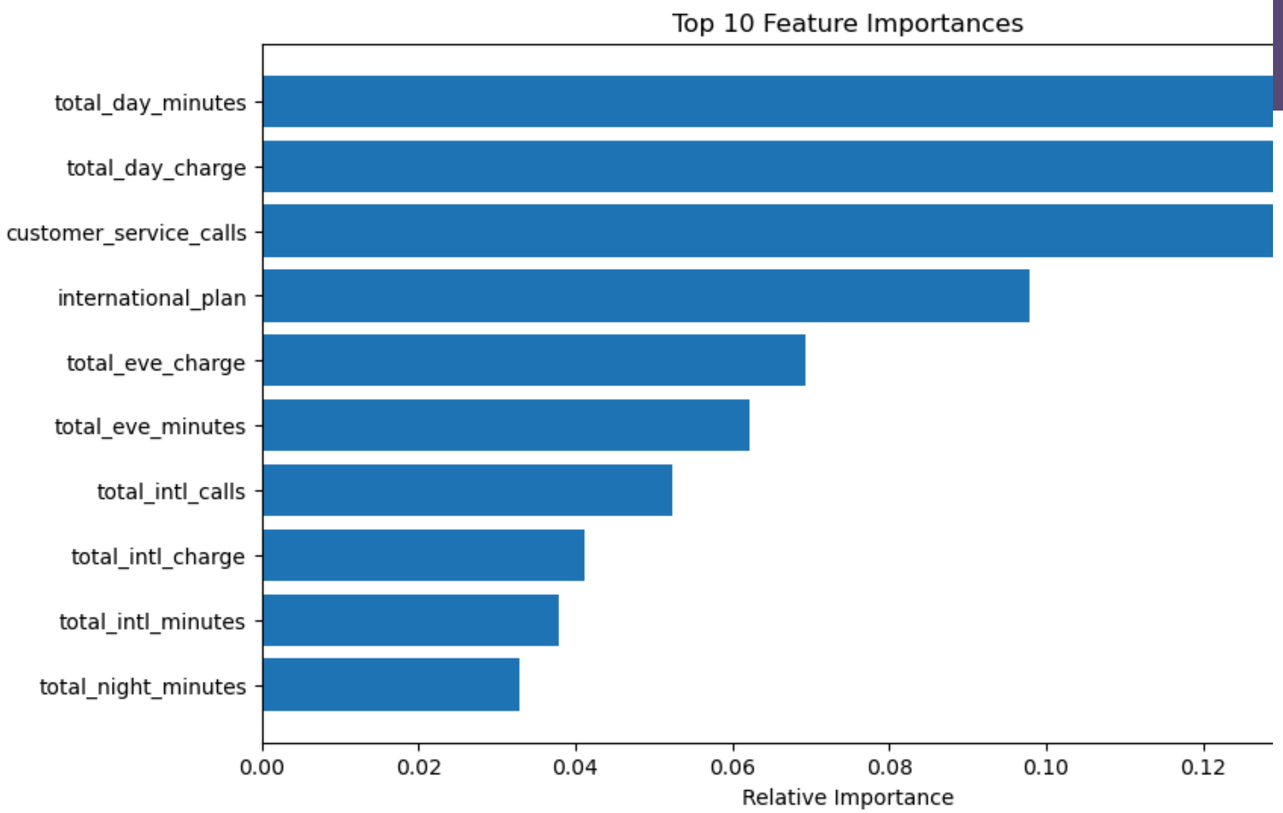
MODELING

This being a classification problem ,various classification models were put in practice.

They include :

- Decision Tree
- Logistic Regression
- Random Forest
- XGBoost

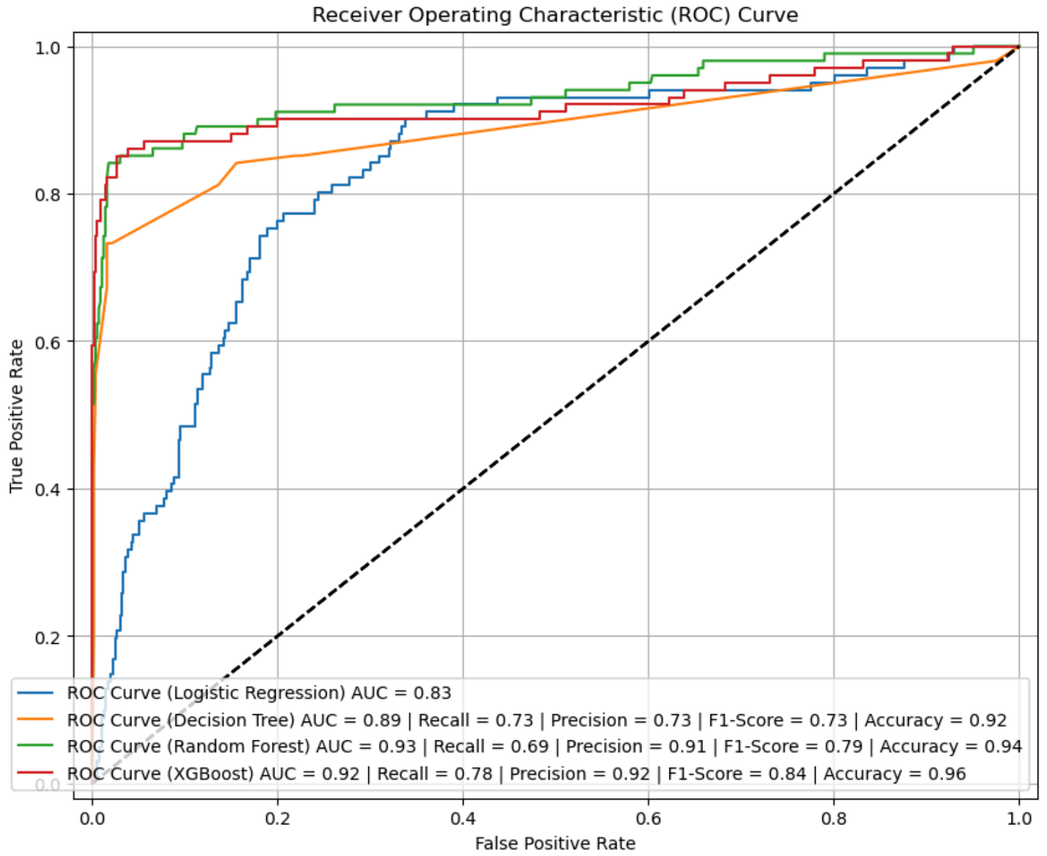
MODEL EVALUATION



Total day minutes,total day charge and customer service charge have a high effect on the customer churn rates.

MODEL EVALUATION

- **Logistic Regression: AUC = 0.83**
- **Decision Tree: AUC = 0.89**
- **Random Forest: AUC = 0.93**
- **XGBoost: AUC = 0.92**



Recommendations

- **Focus on improving customer service**
- **Evaluate the current international plan's competitiveness in terms of price and features**
- **better day-time packages or discounts to heavy users to make them feel valued and reduce their churn rates.**
-

Conclussions

XGBoost is the best algorithm to use due to its highest recall and F1-Score, making it effective at identifying churn while minimizing false positives. If computational resources are limited or model interpretability is important, Random Forest is overall a reliable choice because of its fully-superior ability to predict customer churns and also its not so much limited computational resources.