

# SYRIATEL - CUSTOMER-CHURN- PREDICTION

By Samuel Gichanga



# BUSINESS UNDERSTANDING

## Problem Statement

- SyriaTel, a telecommunications company, is facing a high churn rate, with many customers discontinuing their services and switching to competitors. The company wants to address this issue by developing a customer churn prediction model. By analyzing the dataset, SyriaTel aims to gain insights into factors associated with churn, with the goal of reducing churn rate, increasing customer retention, and improving overall profitability.



Identify the factors that are most likely to lead to customer churn.

Develop a model that can accurately predict which customers are at risk of churning.

Take proactive steps to retain customers who are at risk of churning.



## SPECIFIC OBJECTIVES

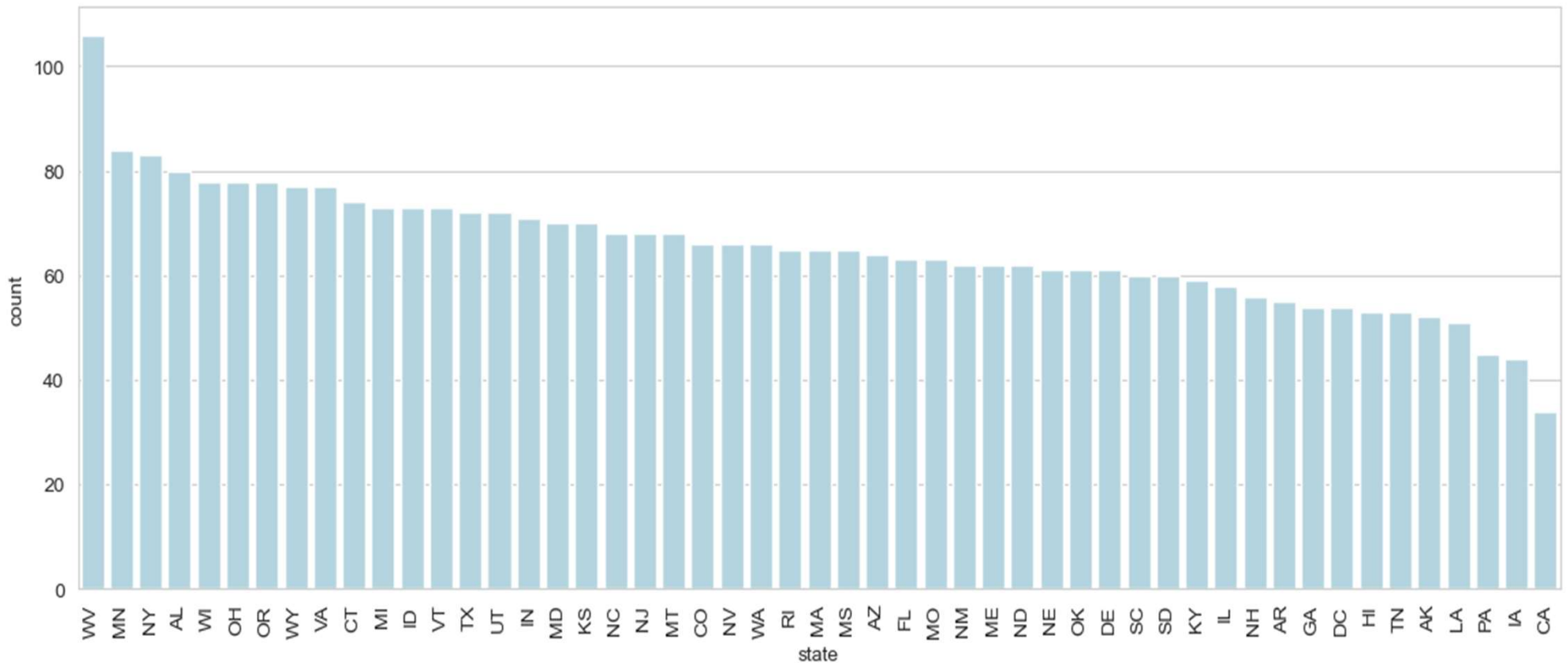
# DATA UNDERSTANDING.

The dataset provided information on the following features for each customer:

- \* US State
- \* Length of account
- \* Area code + phone numbers
- \* If the customer has an international plan
- \* If the customer has a voicemail plan
- \* No. of voicemail messages
- \* Breakdown of call minutes for day, evening, night, and international
- \* Breakdown of call charges for day, evening, night, and international
- \* Breakdown of no. of calls for day, evening, night, and international
- \* No. of calls to customer service
- \* If they have churned

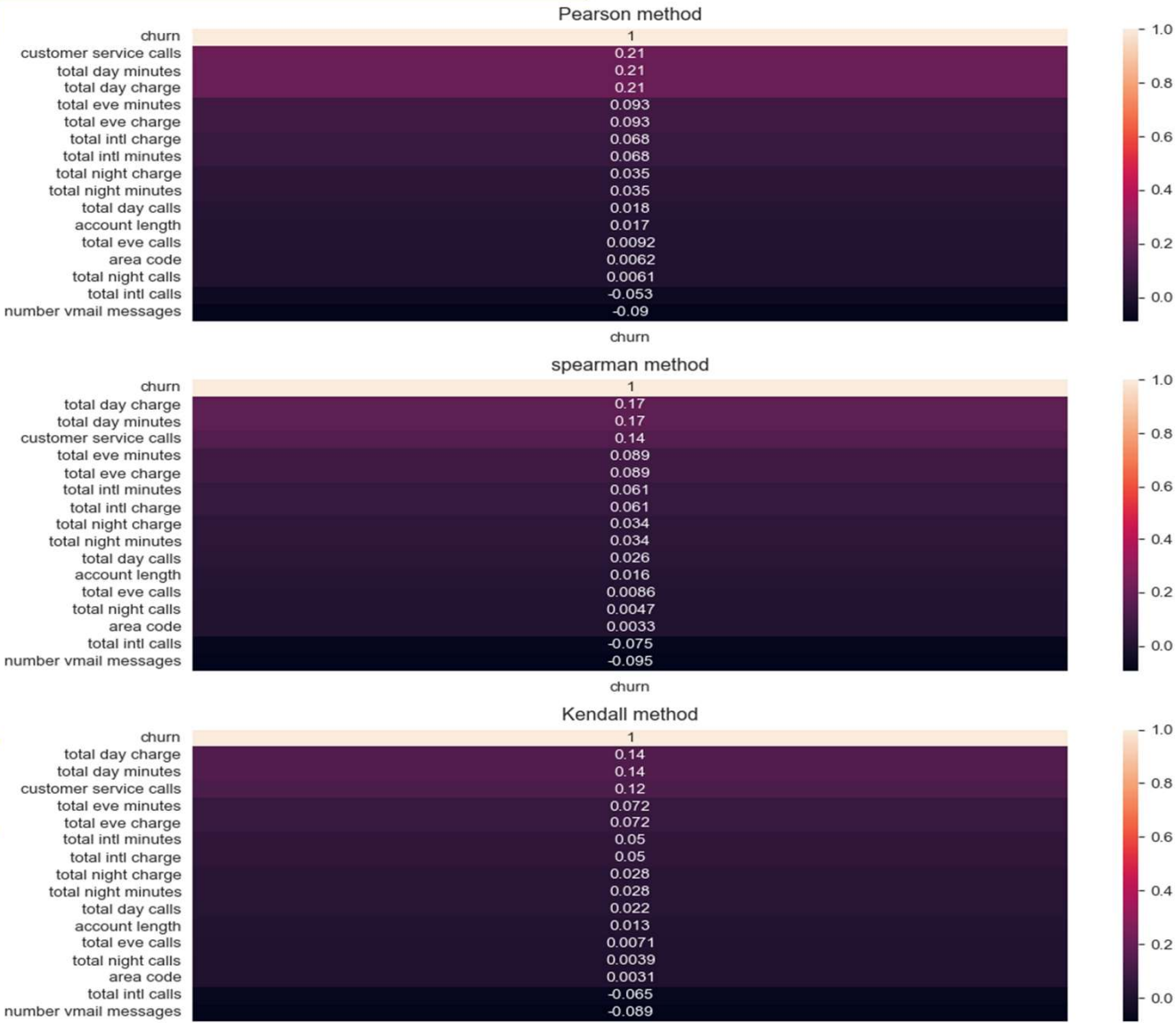
# NO. OF CHURNS PER STATE

From the visualization of the features I realized a trend in the certain states churning more than others, so I plotted the no. of churns per state below.



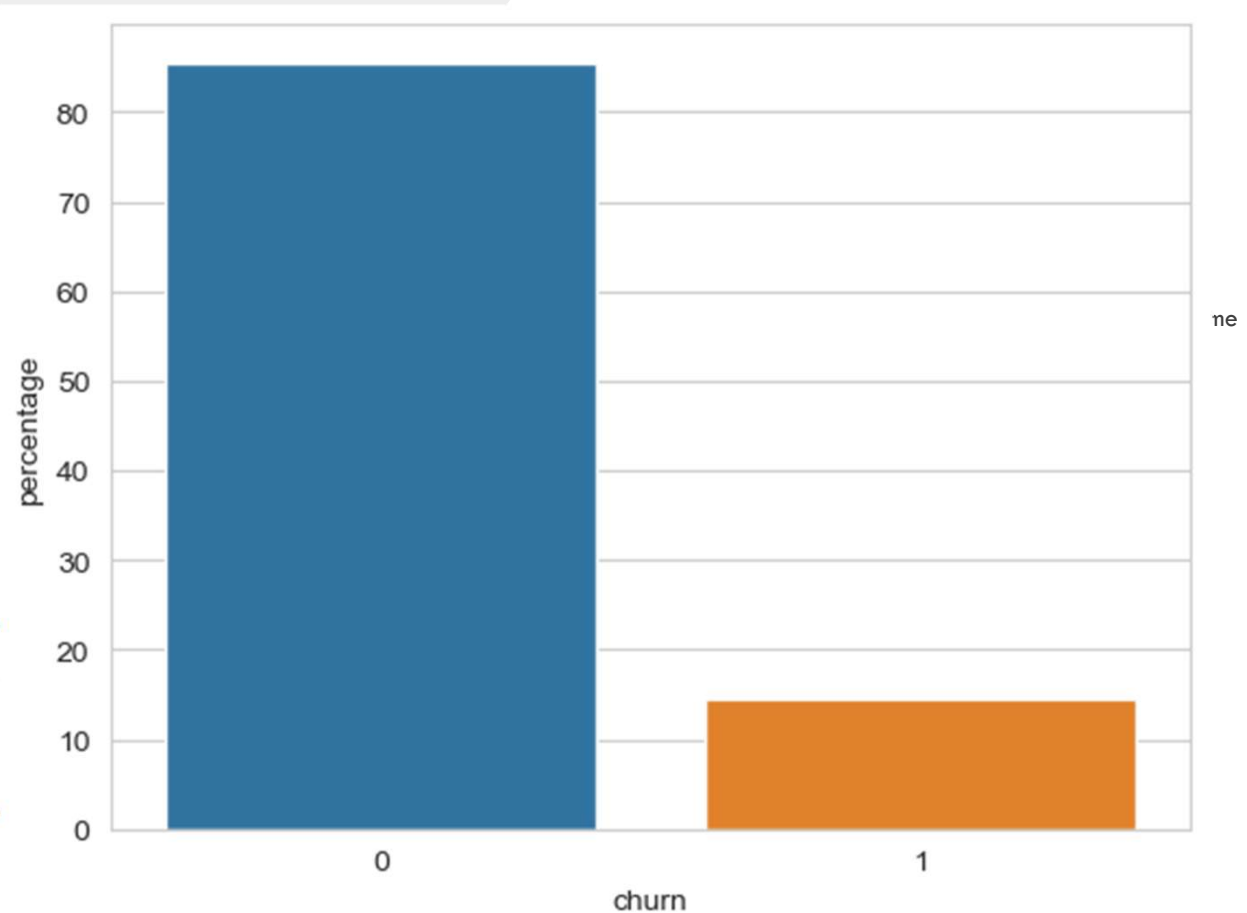
# CORRELATION

From the plots is evident that all numerical features shows only little correlation with target variable, with total init calls and number of vmail messages having weak negative correlation and all the others show weak positive correlation with the target variable.



# CHURN PERCENTAGE

Of the 3,333 customers in the dataset, 483 have terminated their contract. That is 14.5% of customers lost. The distribution of the binary classes shows a data imbalance. This needs to be addressed before modeling as an unbalanced feature can cause the model to make false predictions.





# MODELING

Since the goal is to reduce the number of customer churns, I want to reduce the number of number of customers that leave since they were not identified as a churn risk, or simply, the number false negatives. This means that we want to prioritize recall, but not make the model so inaccurate that all customers are identified as potential churners.

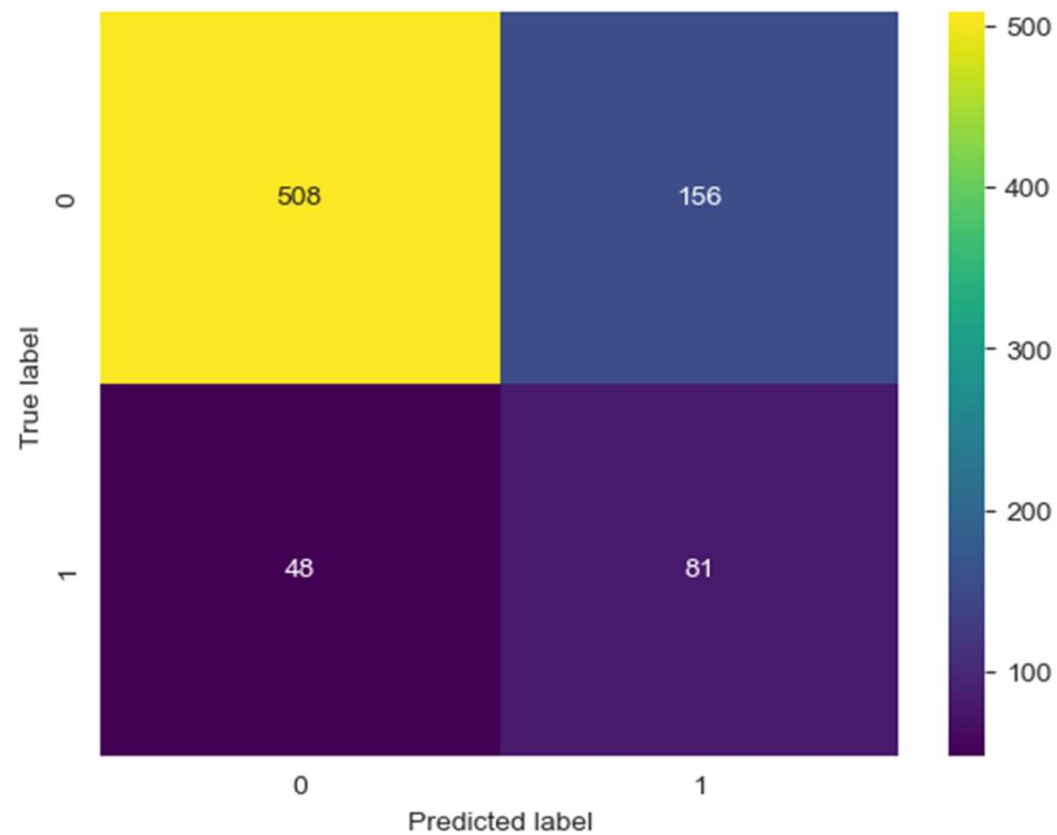
I built a model that can predict the customer churn based on the features in our dataset. The model will be evaluated on the recall score. Specifically, if it achieves an recall score of 80% or higher, it will be considered a success.

In order to achieve the targets stipulated in the project proposal, we will be using the following algorithms:

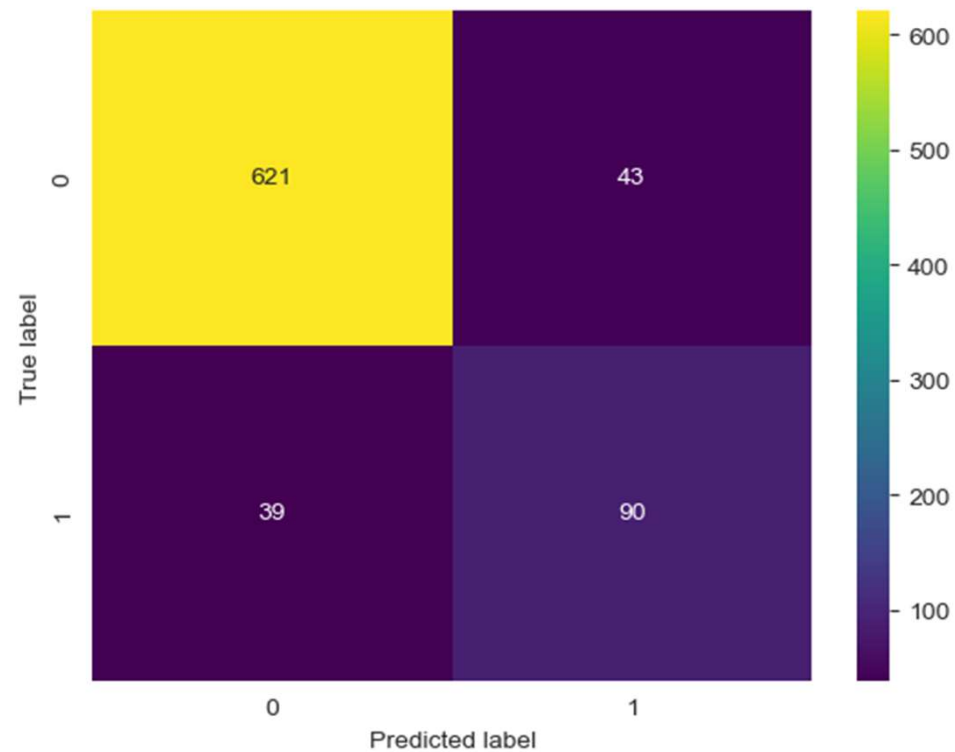
1. Logistic Regression
2. Decision Tree
3. Random Forest
4. XG Boost



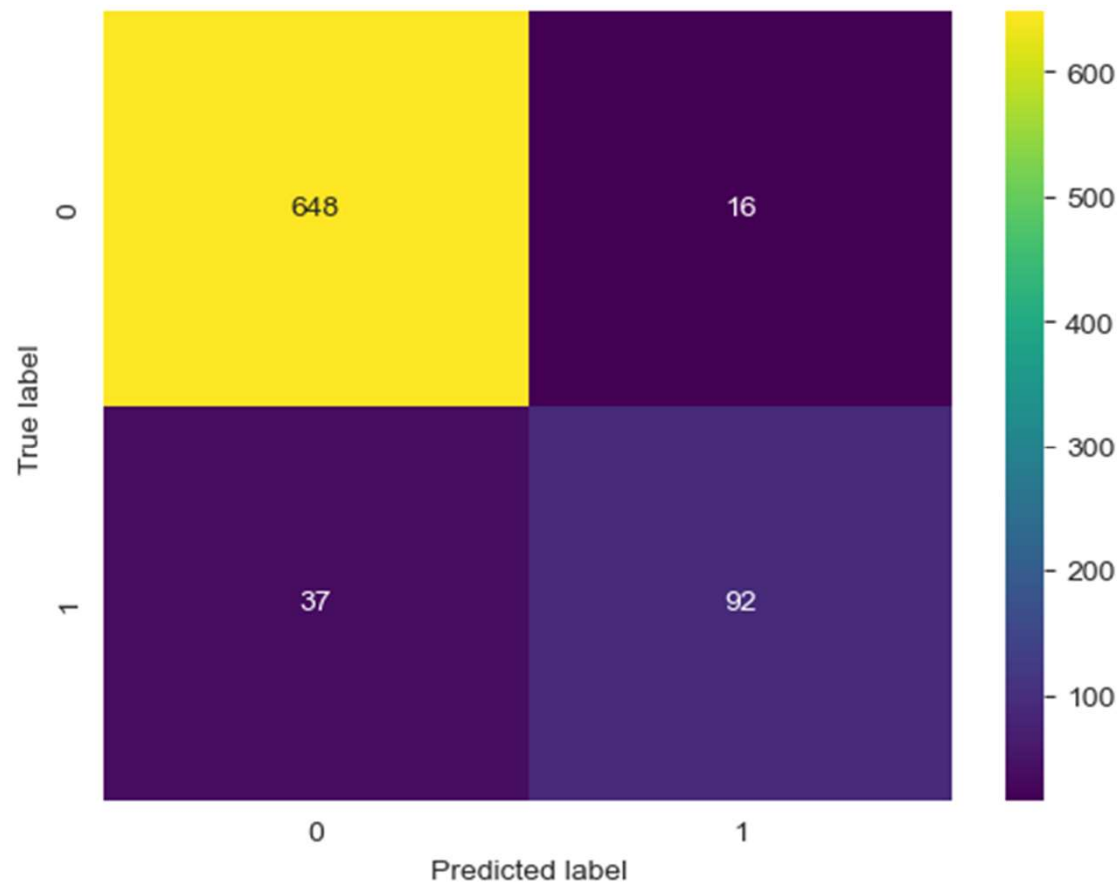
## MODEL 1: LOGISTIC REGRESSION (CONFUSION MATRIX)



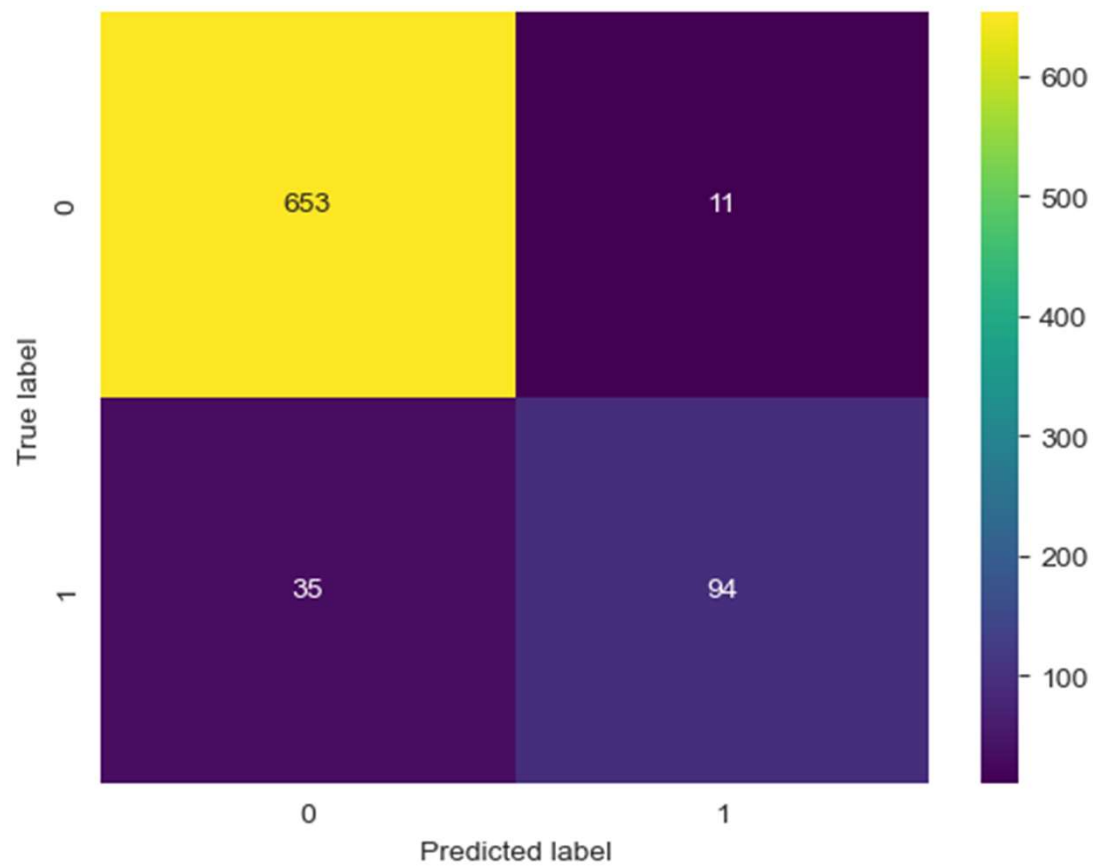
## MODEL 2: DECISION TREES (CONFUSION MATRIX)



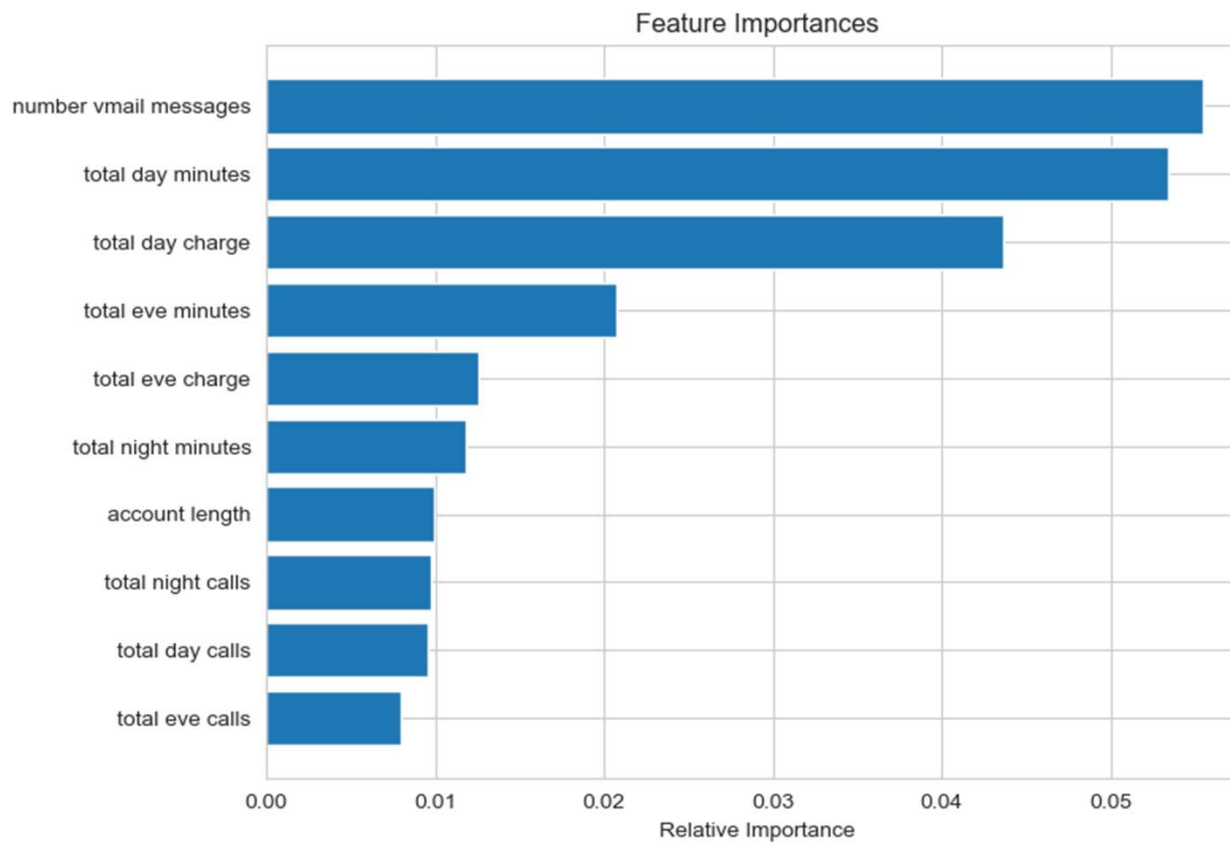
## MODEL 3: RANDOM FOREST (CONFUSION MATRIX)



## MODEL 4: XGBOOST (CONFUSION MATRIX)



# TOP 10 FEATURE IMPORTANCE GRAPH





# FINDINGS

Within the dataset, it was evident that the SyriaTel business model was to charge customers based on the number of minutes that they used. However, it is evident that within the contributors of high customer churn, all of the factors are leading to a higher bill that is deterring the customer from continuing their phone plan.



# RECOMMEN DATIONS

Evaluate the pricing structure for day. Consider adjusting pricing plans or introducing discounted packages to address the higher charges associated with customers who churn.

Since Most of the customers are from West Virginia, Minnesota, NewYork, Alabama and Wisconsin, they should Offer discounts or promotional offers to the customers from this states. as these areas have a higher churn rate. This can help incentivize customers to stay with the company.

Develop customer retention strategies specific to high night time minute users. This could involve offering retention incentives or discounts to keep these customers loyal.

Personalized Offers; Use the information about high daytime minute usage to create personalized offers and incentives for customers who are at risk of churning. Tailor these offers to match their usage patterns and preferences.

**THANK YOU.**