



CUSTOMER CHURN PREDICTION

MAUREEN KITANG'A

OVERVIEW



- Customer churn is a significant challenge for businesses, particularly in the telecom industry.
- Understanding the factors that contribute to customer churn and being able to predict it can help companies develop strategies to retain customers and reduce revenue losses.

PROJECT WORKFLOW



Business
Understanding



Data
Understanding



Modeling



Evaluation

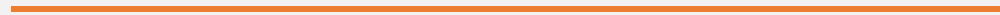


Conclusion

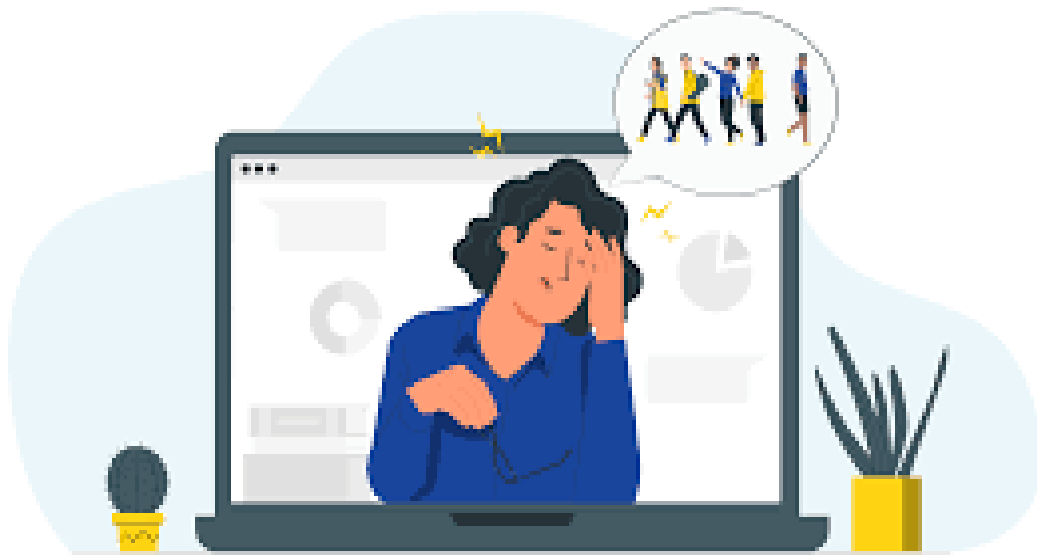


Recommendation

BUSINESS UNDERSTANDING



BUSINESS PROBLEM



Syriatel, a telecommunications company, is facing a high churn rate, with many customers discontinuing their services and switching to competitors.



PROJECT OBJECTIVES

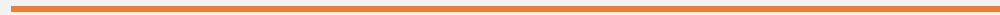
- Identify the factors that are most likely to lead to customer churn.
- Develop a model that can accurately predict which customers are at risk of churning.
- Achieve a recall score of at least 80%.

DATA UNDERSTANDING

- We will work with a customer churn dataset from the telecom industry sourced from Kaggle website.
- It contains 3,333 records and 21 columns(4 categorical, 17 numerical).

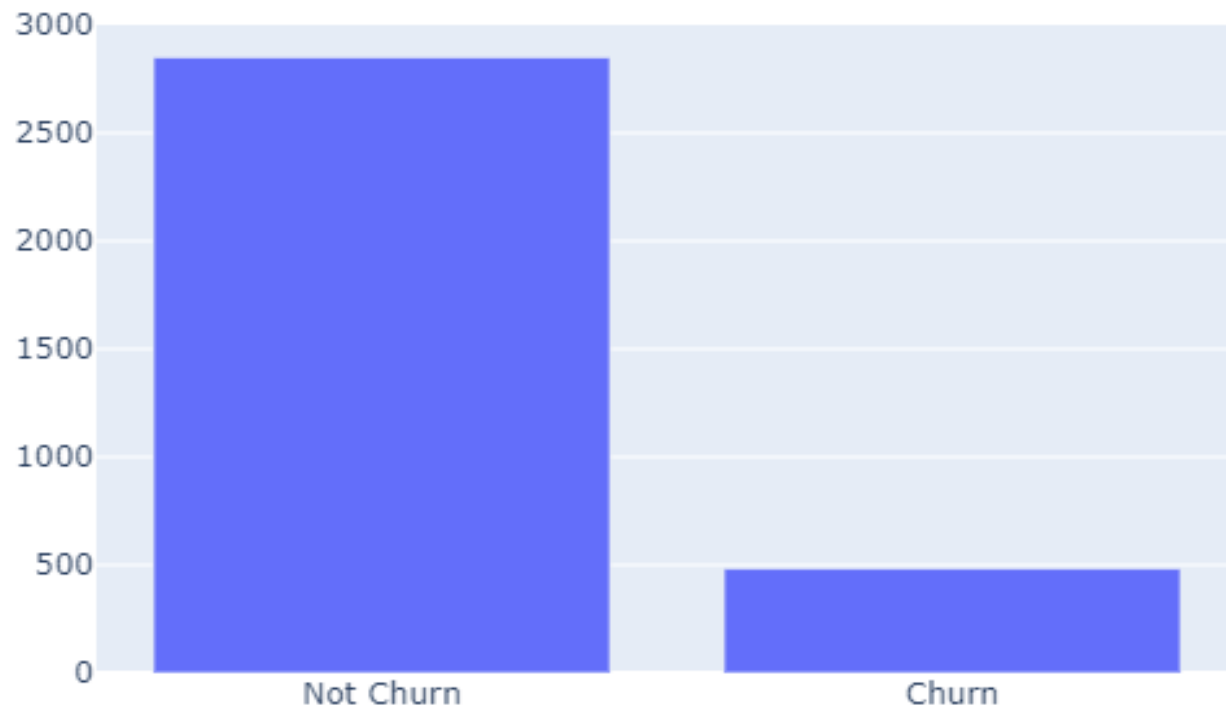


ANALYSIS RESULTS...



How many customers have churned?

Churn Distribution



Total number of
customers;

3,333

Number of customers
that churned;

483

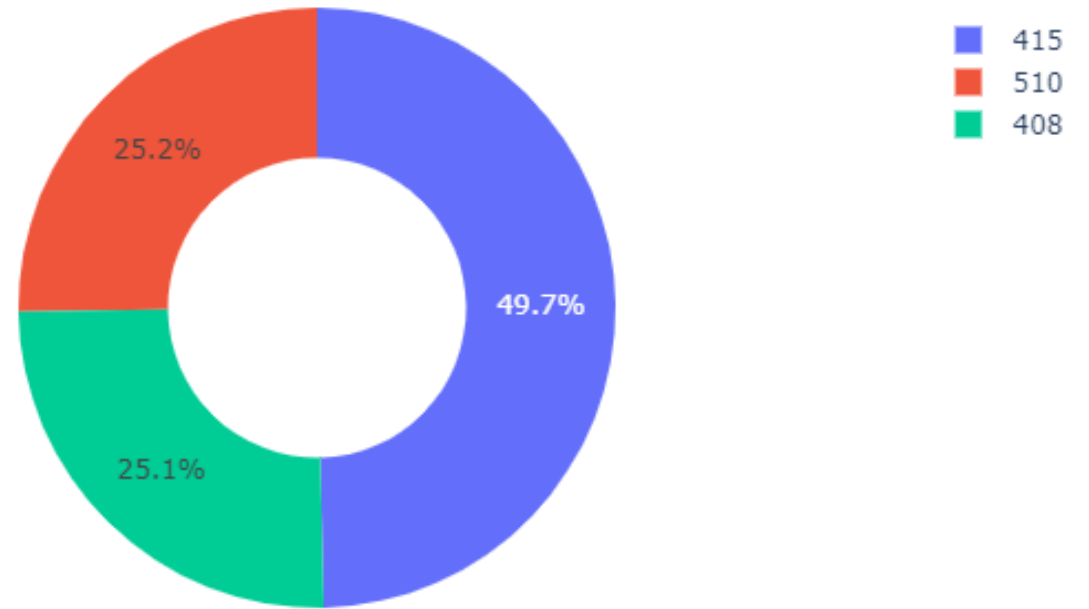
Churn Percentage;

14.5%

Area Code Data Distribution

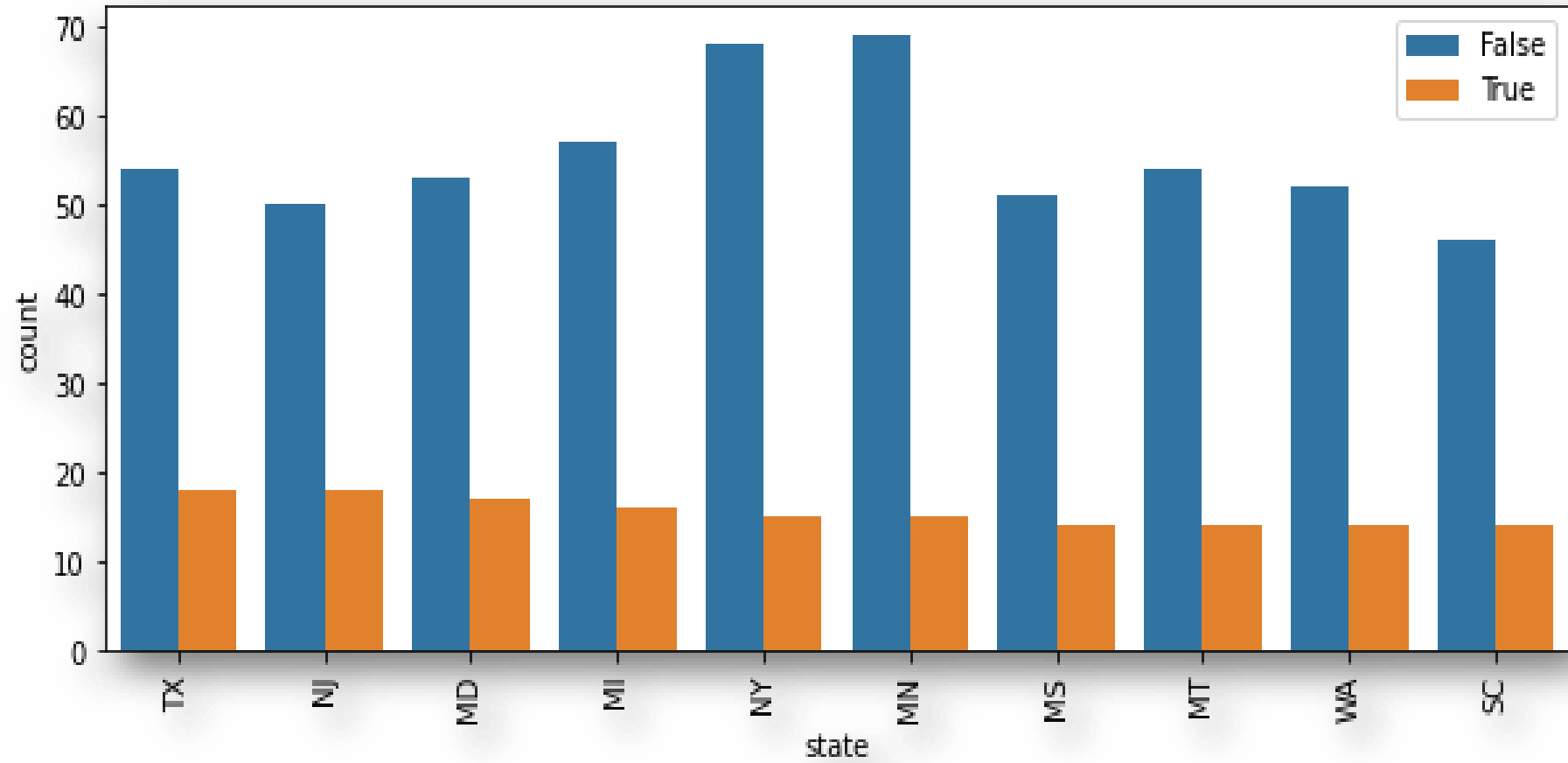
The area code 415 has the highest proportion, accounting for (49.7%) of the total customer base.

Both area codes 510 and 408 have an equal number of customers.(25.2% and 25.1%) respectively.

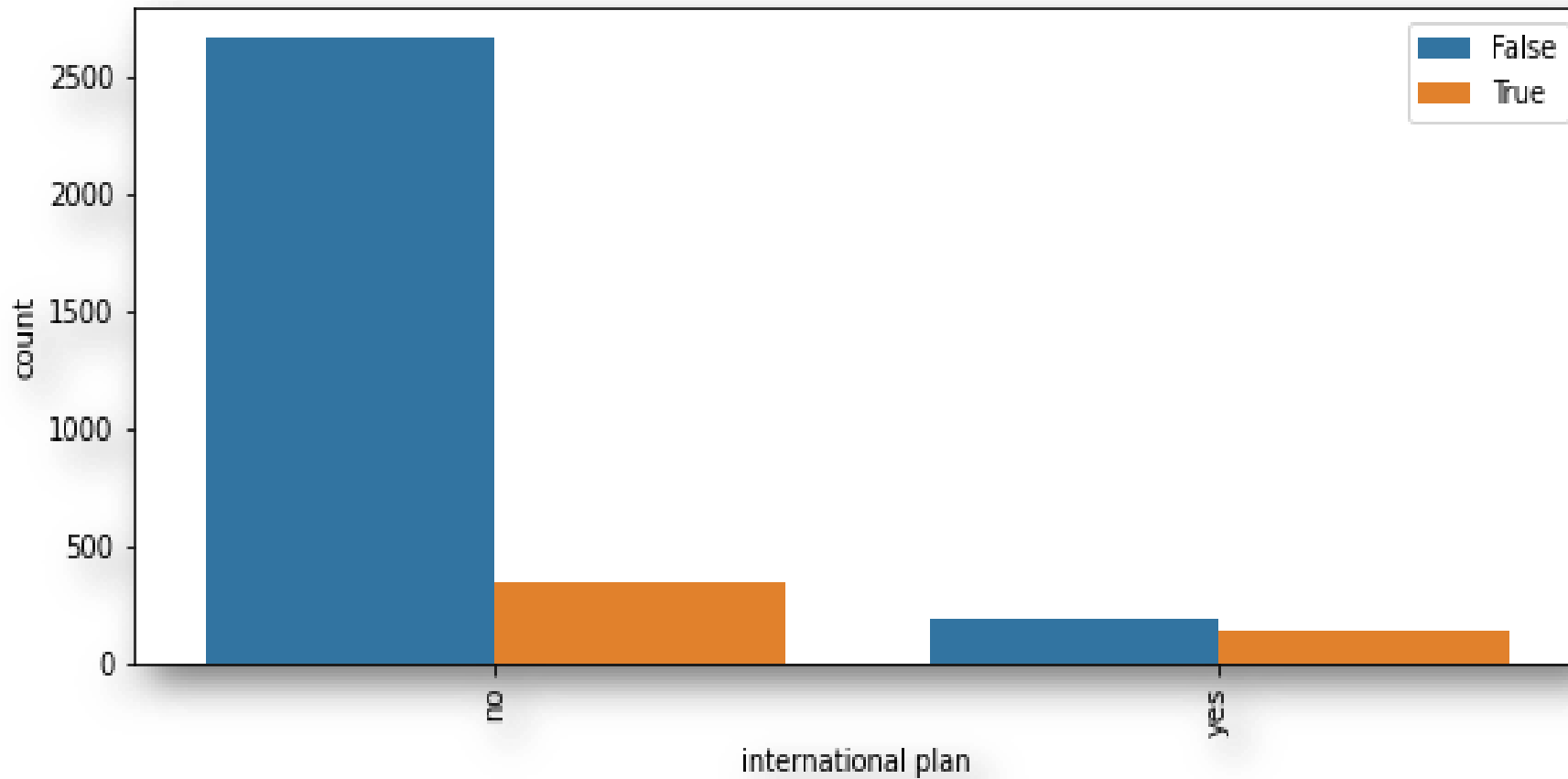


Which states have the highest churn rate?

Texas and New Jersey are leading followed by Maryland, Miami and New York.

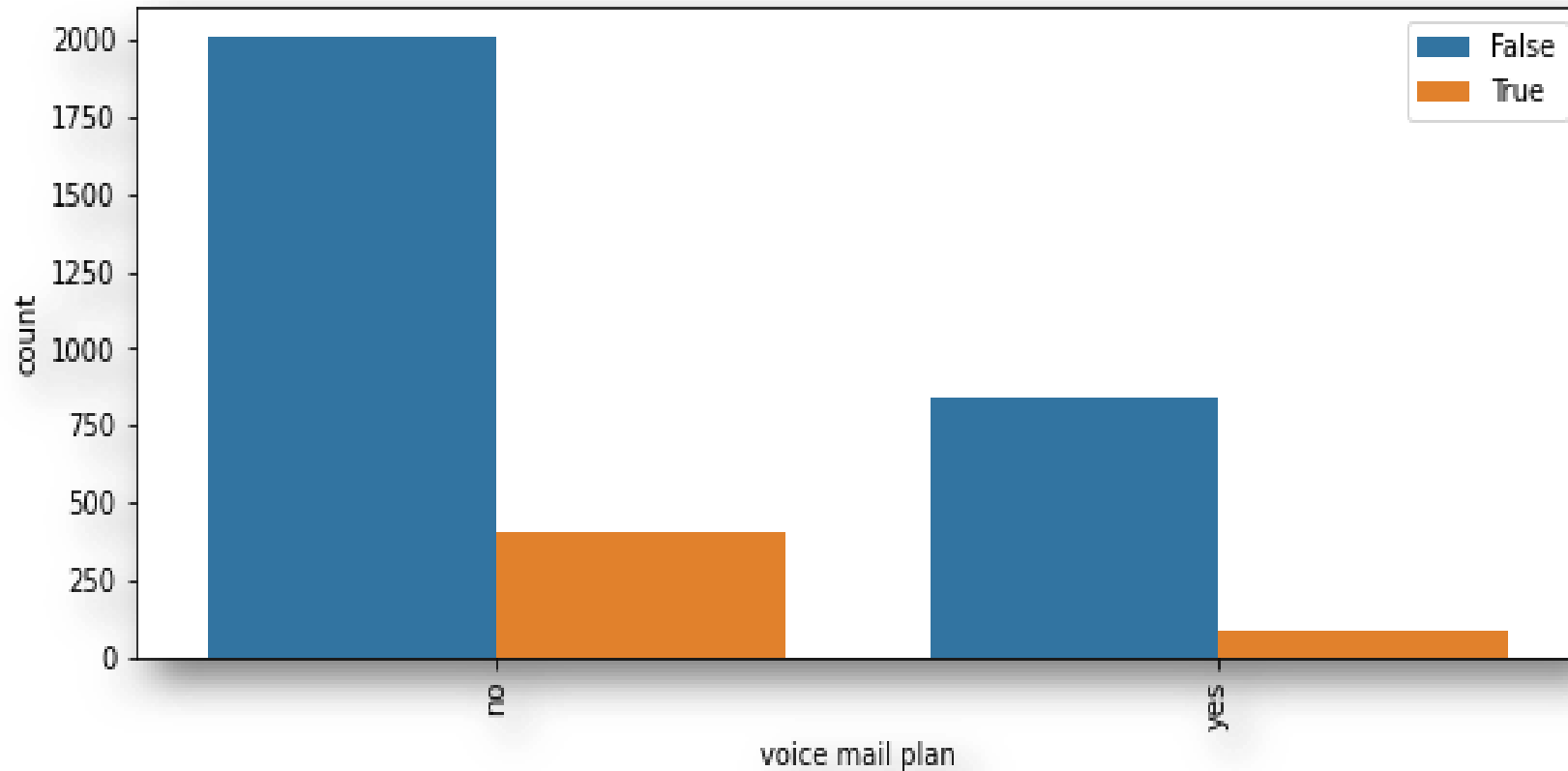


How does International plan impact Churn?



Of the 483 customers that churned, majority did not have an international plan.

How does Voicemail plan affect Churn?



Of the 483 customers that churned, majority did not have a voice mail plan.

MODELING



Algorithms Used

- Logistic Regression
- Decision Tree
- Random Forest
- XG Boost

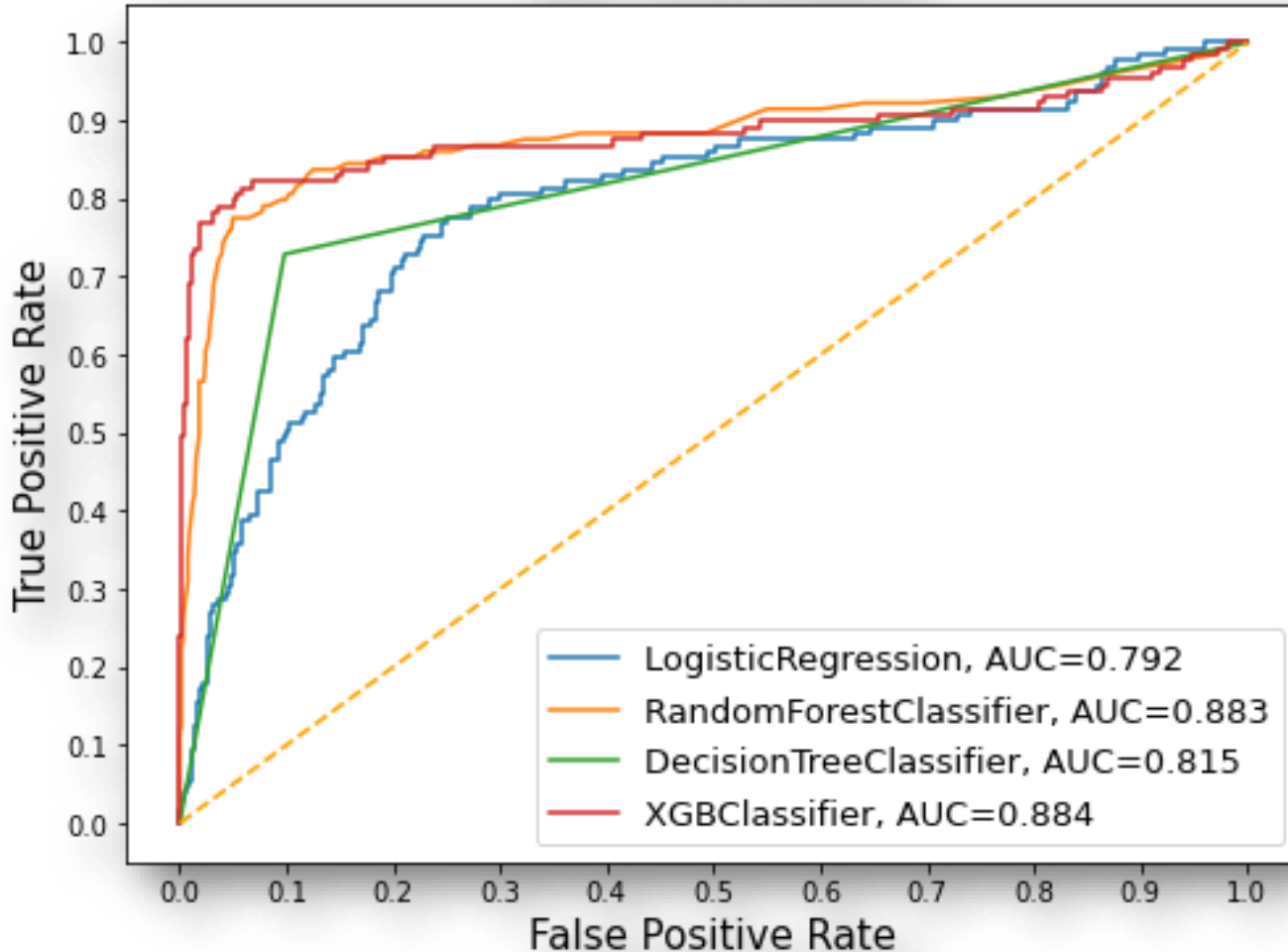
Metrics Used

- Recall-focuses on the model's ability to correctly identify positive cases.
- ROC_AUC Curve-plots the true positive rate (sensitivity) against the false positive rate ($1 - \text{specificity}$) at various classification thresholds.



EVALUATION

ROC Curve Analysis



Based on recall score;

- XGBoost - 0.77
- Random Forest - 0.74
- Logistic Regression - 0.74
- Decision Tree - 0.73

Based on ROC_AUC curve;

- XGBoost - 0.884
- Random Forest - 0.883
- Decision Tree - 0.815
- Logistic Regression - 0.792

MODEL TUNING-RESULTS

- The tuned Random Forest model achieved a recall score of 0.76 and an ROC_AUC of 0.89, which was an improvement from the untuned model.
- The tuned XGBoost model also achieved an improvement, with a recall score of 0.79 and an ROC_AUC of 0.89.

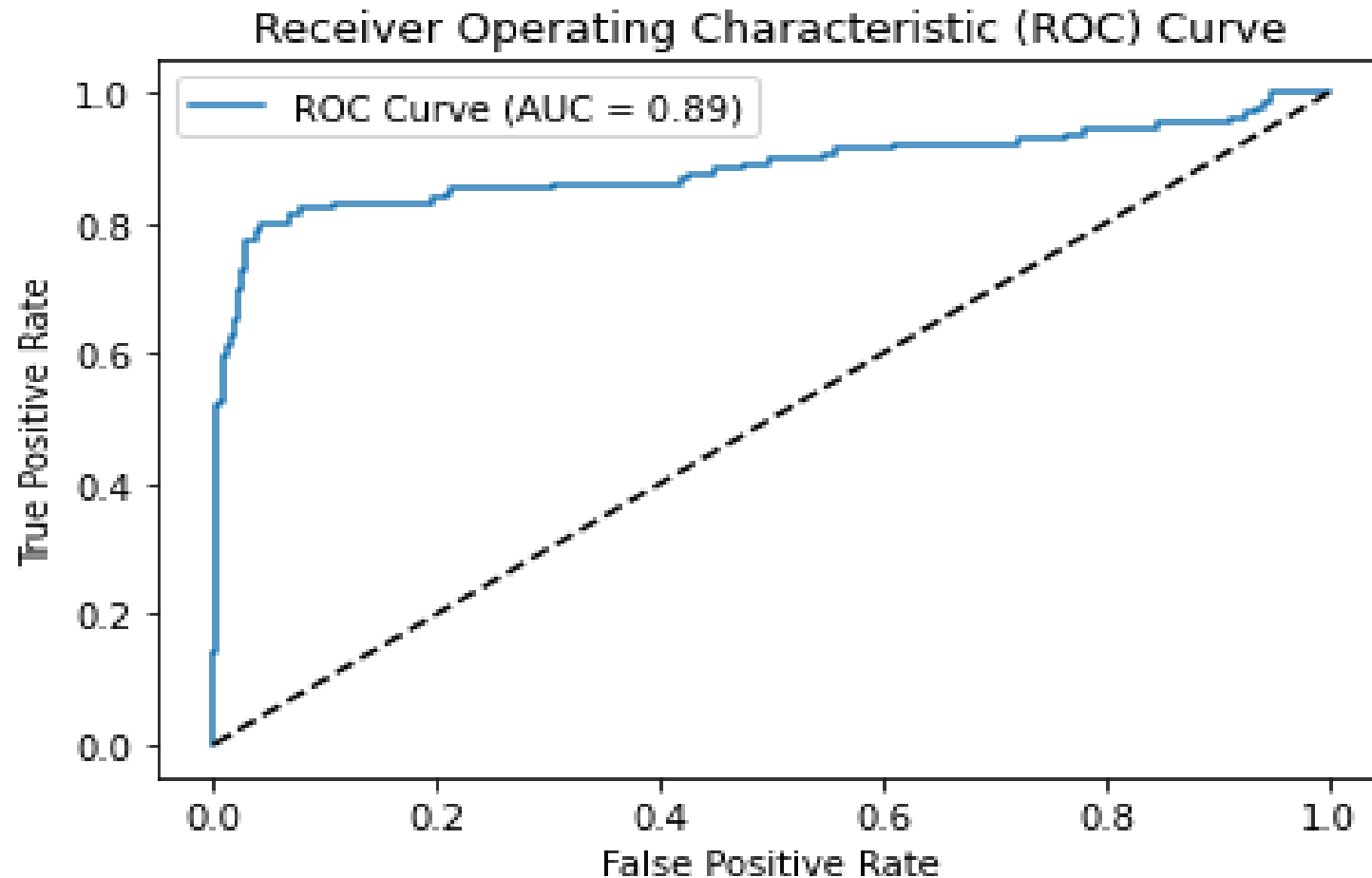
XGBoost Model

Recall score

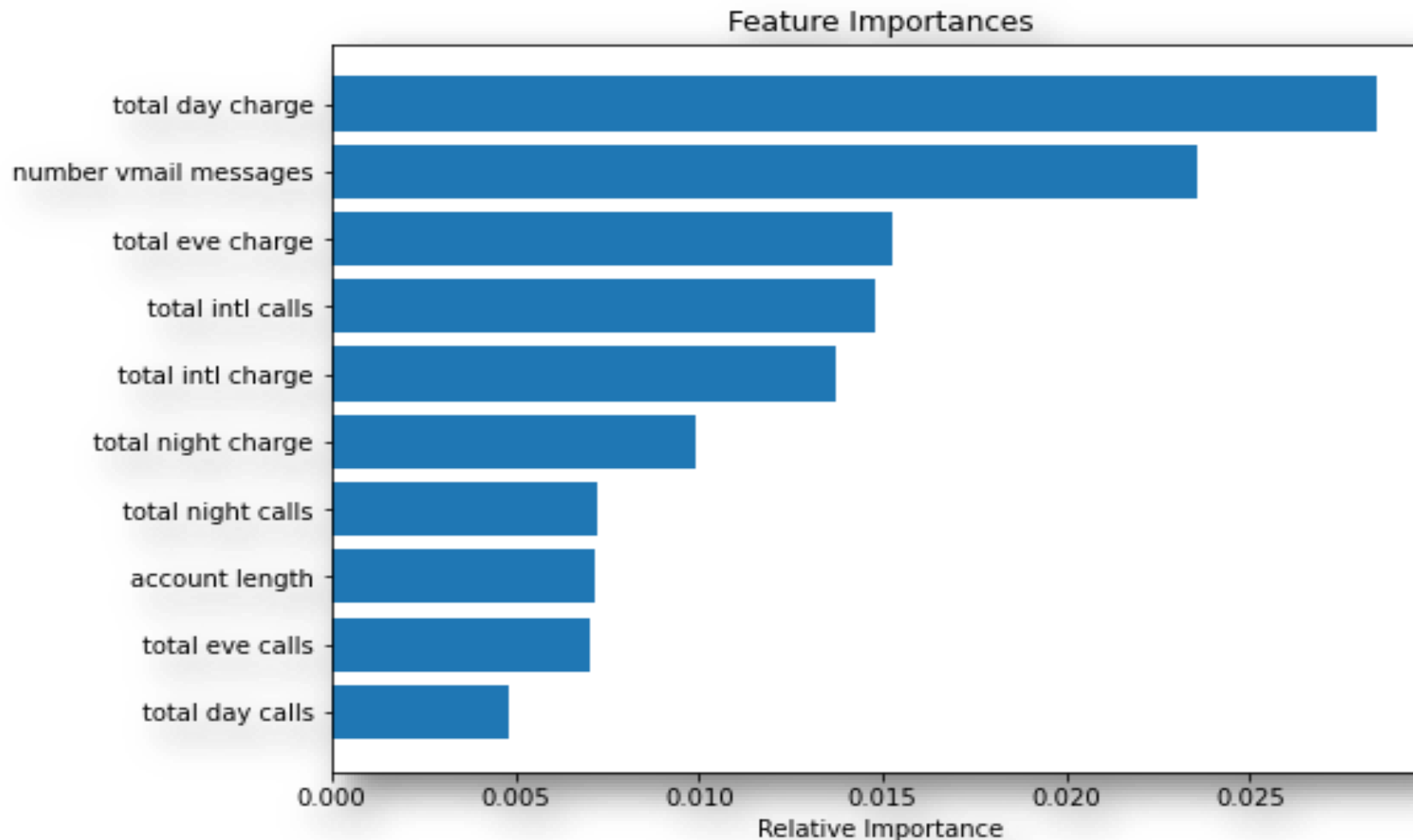
0.79

ROC_AUC score

0.89



Which features impacted churn according to our model?



- **Total day charge**; customers who had more charges during the day churned the most.
- **number vmail messages**; customers that received fewer number of voicemail messages churned the most
- **total eve charge**; customers who had more charges during evening time churned the most.

CONCLUSION

- The recall score of our XGB classifier was 79%. While this is still a good predictive model, we would like to undertake further feature engineering to boost this recall score if we had more time.
- total day charge, number vmail charge, total eve charge are the most important features in determining customer churn.
- We achieved our objectives to be able to predict customer churn and had an acceptable recall score.

RECOMMENDATIONS

- Offer discounts or promotional offers to customers in area code 415 and 510, as these areas have a higher churn rate.
- Improve customer service quality and reduce the number of customer service calls. Enhance training programs for customer service representatives to ensure prompt and effective resolution of customer issues, leading to higher customer satisfaction and reduced churn.
- Evaluate the pricing structure for day, evening, night, and international charges. Consider adjusting pricing plans or introducing discounted packages to address the higher charges associated with customers who churn.

RECOMMENDATIONS

- Focus on customer retention strategies in states with higher churn rates, such as Texas, New Jersey, Maryland, Miami, and New York. This can involve targeted marketing campaigns, personalized offers, or improved customer support tailored to the specific needs and preferences of customers in those states.
- Enhance the value proposition of the voicemail plan to increase adoption among customers. Highlight the benefits and convenience of voicemail services, and consider offering additional features or discounts to encourage customers to sign up.



THANK YOU:)

For more information, reach out
ndungek66@gmail.com