



SyriaTel Communications

**Predicting &
Preventing Customer
Churn**





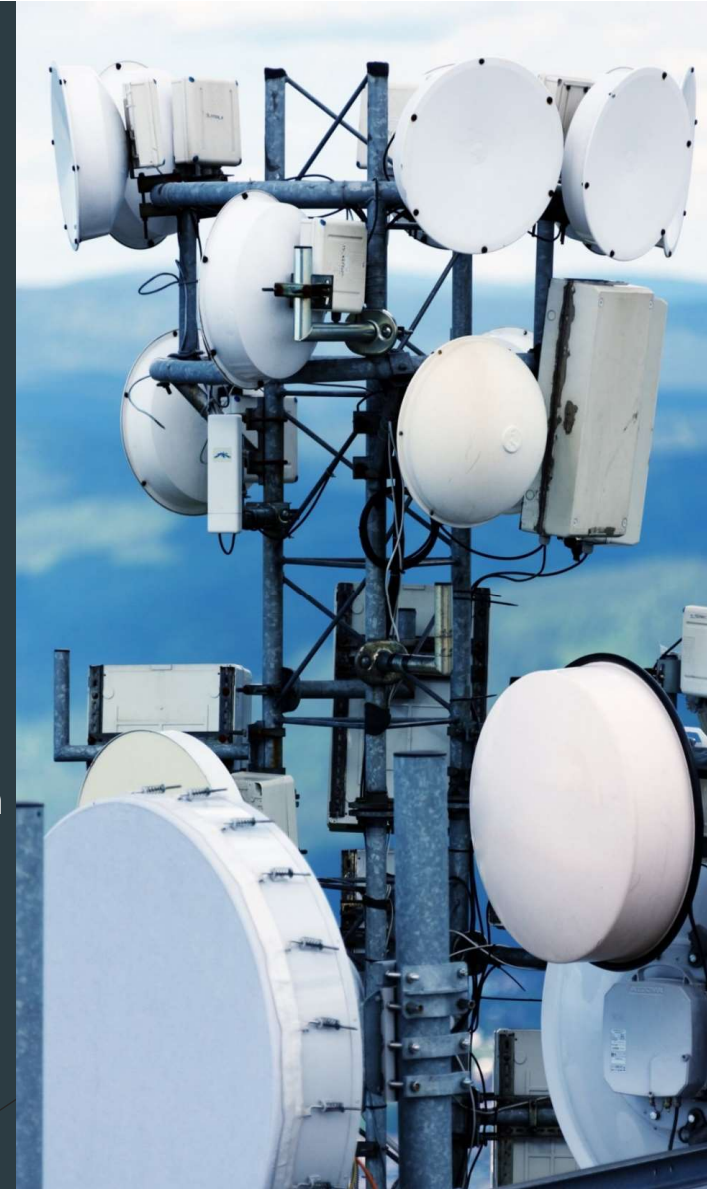
INTRODUCTION

In the ever-evolving landscape of the telecommunications industry, SyriaTel, a key player, confronts the critical challenge of customer churn—a phenomenon where subscribers discontinue their association with the provider. Against the backdrop of rapid technological advancements, shifting consumer preferences, and intense market competition, the need to understand and predict customer behavior is more pressing than ever. This project emerges as a response to this imperative, aspiring to construct a sophisticated binary classification model that not only anticipates customer churn but provides actionable insights for SyriaTel's strategic approach to customer retention.



Problem statement

In the realm of telecommunications, customer churn remains a critical challenge, and SyriaTel, a prominent player in the industry, is not immune to its impact. The dataset at hand encapsulates a diverse array of customer attributes, usage patterns, and service interactions, presenting a complex scenario for predicting customer churn accurately. The multifaceted challenges inherent in the dataset include data variability, non-normality of residuals, model evaluation complexities, outliers, data imbalance, categorical data intricacies, missing data, and multicollinearity.



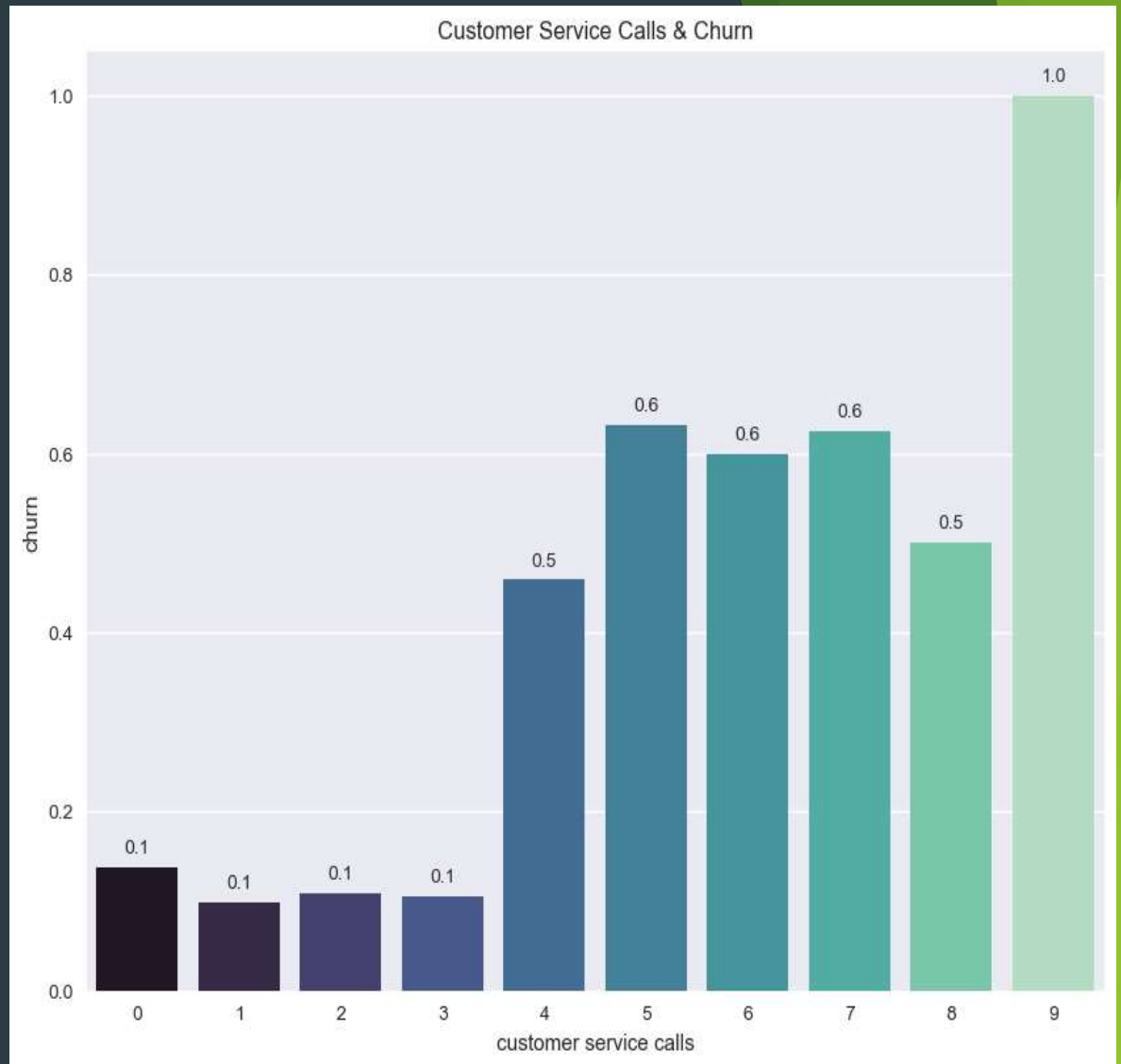
OBJECTIVES

- To construct a predictive model for customer churn that takes into account geographical variations, providing insights into regional variations in churn likelihood.
- ★ To explore usage metrics within the SyriaTel dataset and integrate them into the churn prediction model, enhancing its predictive accuracy by considering customer usage patterns.
- To analyze customer service call data and extract relevant features, incorporating them into the churn prediction model.
- ★ To develop a predictive model that anticipates customer churn for SyriaTel, proactively identifying potential churners before they discontinue their association with the telecom company.
- To select suitable machine learning algorithms for the churn prediction model and train them on the preprocessed dataset.
- ★ To rigorously evaluate the performance of the developed predictive model, emphasizing interpretability.

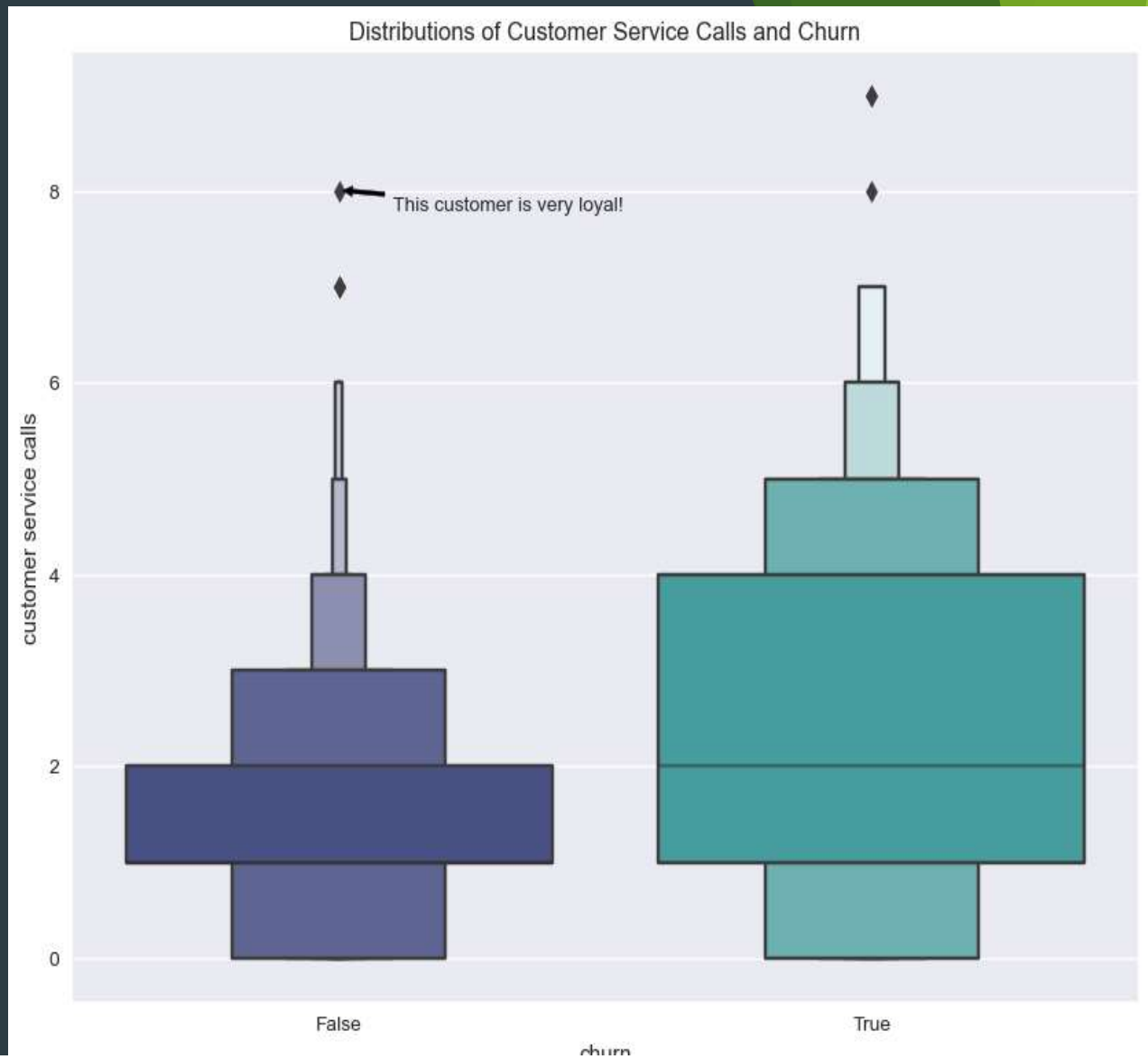
Analysis

Calls to Customer Service

An increase in churn after 3 calls to customer service



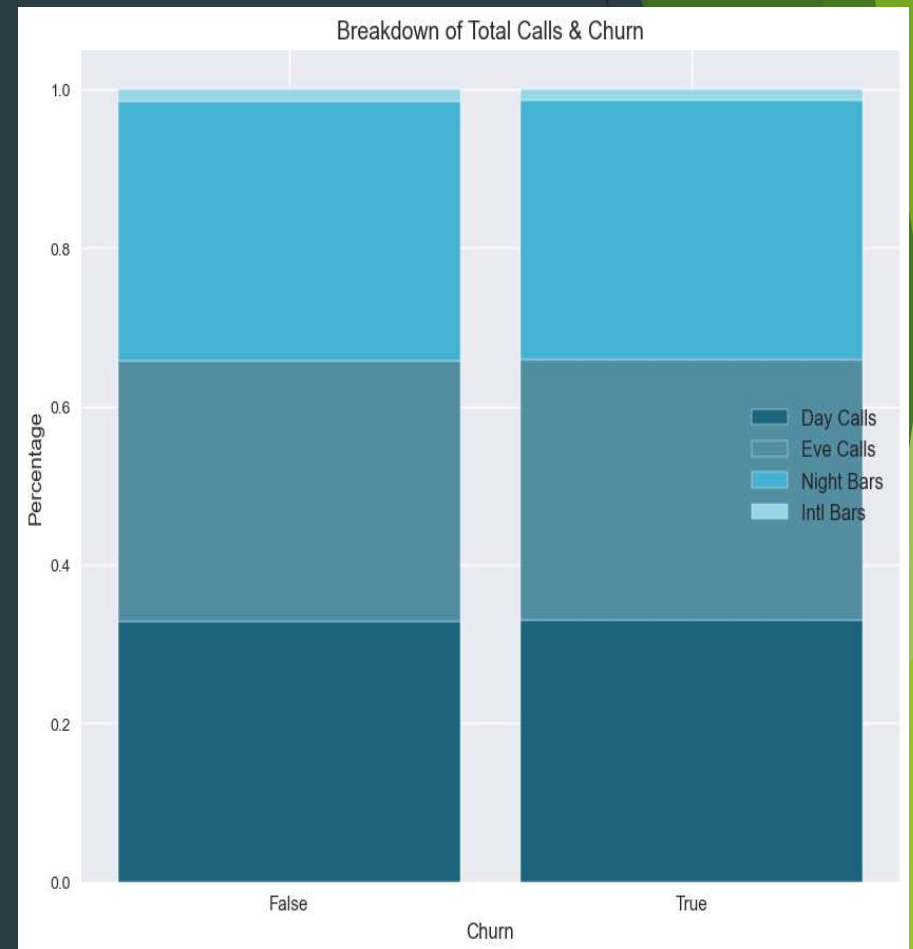
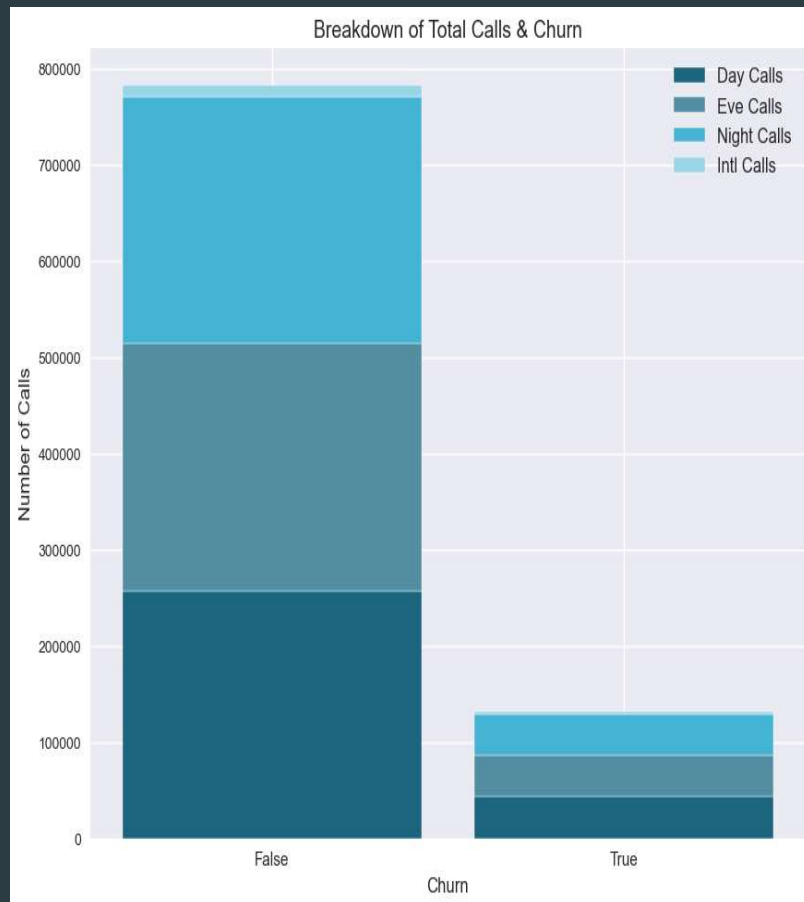
- The majority of customers who did not churn placed 1- 2 calls
- The majority of customers who DID churn placed 1-4 calls



Higher percentage of churn with international plan customers

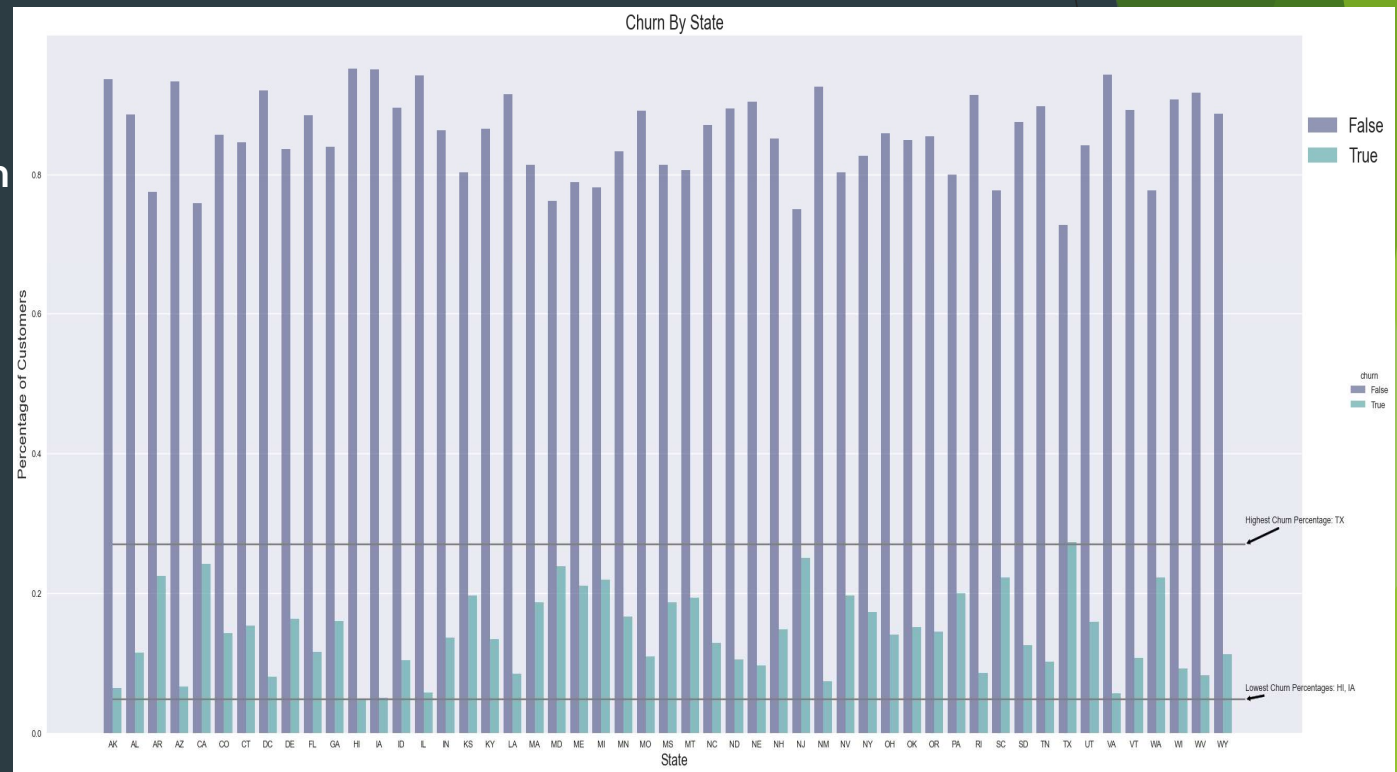


There is virtually no difference between churn and the percentage of day/eve/night/intl



Churn & Customer Location

- Highest churn in Texas
- Lowest churn in Hawaii and Iowa



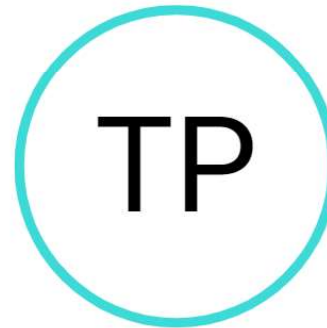
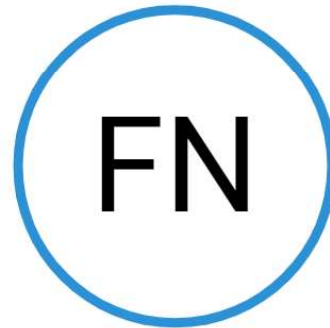
Model	Training Recall	TestiRecall
RandomForestClassifier	1.0	0.734
KNeighborsClassifier	0.97909	0.60638
Gradientboostingclassifier	0.755358	0.744680
GaussianNB	0.752221	0.78723
SVC	0.4668	0.58510

The best performing classifiers here were Gradient Boost, and Gaussian Naive Bayes. They both had the lowest False Negatives and the least overfitting.

CONFUSION MATRIX ANALYSIS

FALSE NEGATIVE

We labeled them as 'not going to churn' when they actually churned.

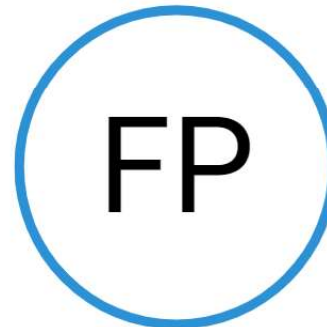
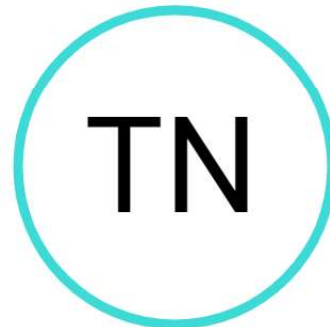


TRUE POSITIVE

We labeled them as 'going to churn' and they did churn.

TRUE NEGATIVE

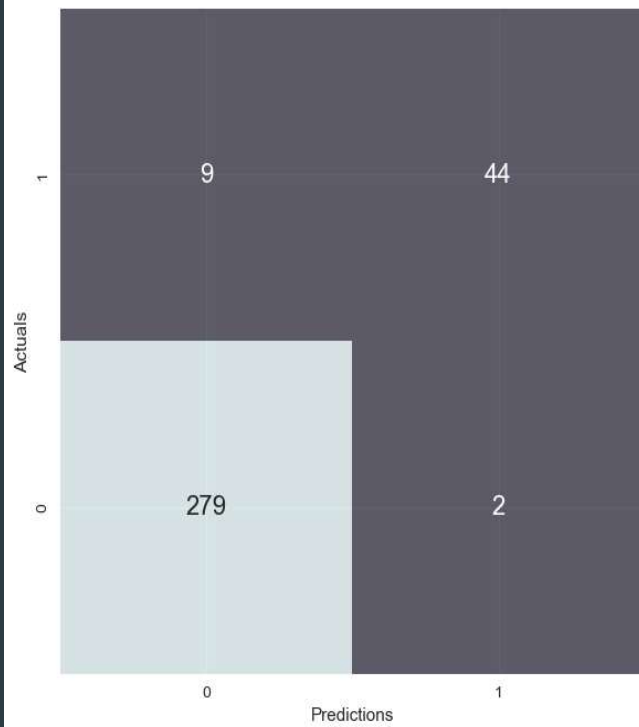
We labeled them as 'not going to churn' and they did not churn.



FALSE POSITIVE

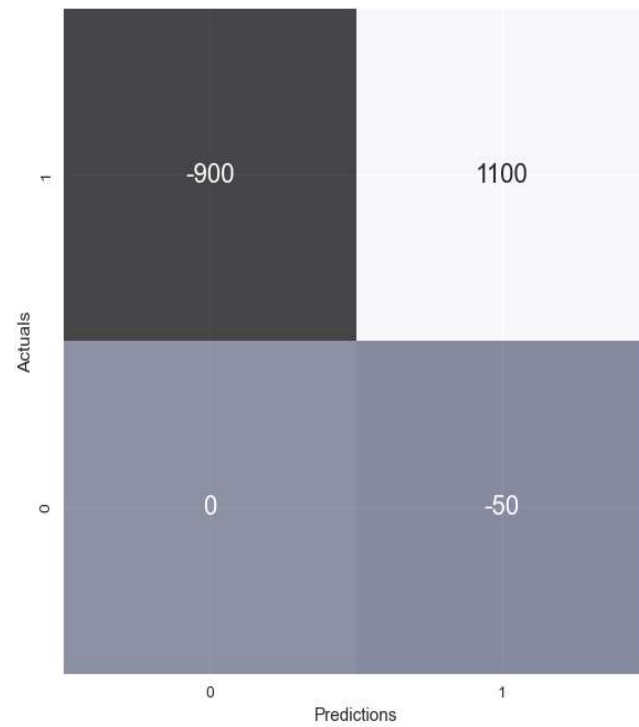
We labeled them as 'going to churn' but they actually stayed.

Validation Set Confusion Matrix



Validation Set Cost Benefit Analysis (\$)

Expected Value: \$0.45 per customer per month



CONCLUSION

We focusing on

- International calls
- Customer service

Recommendation

Obtain additional information on rivals in states with greater churn.

Obtain additional cell signal data across the United States to identify trends in states with greater churn rates.

Examine voicemail logs to determine whether it could be a useful predictor





THANK YOU