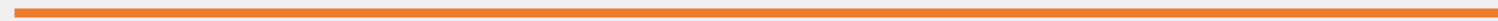


A hand holding a pair of red-handled scissors is positioned to cut a line of small, colorful human figures (red, blue, and green) that are standing on a white surface. The background is a blurred blue and white pattern. In the top right corner, there is a grid of red dots.

Customer Churn Prediction

By Vallary Banda

BUSINESS UNDERSTANDING





Overview

SyriaTel, faces a significant challenge with customer churn, where a significant portion of their customer base is leaving. This not only results in lost revenue but also increases the cost of acquiring new customers to replace those who churn. Given the competitive nature of the telecommunications industry, retaining existing customers is often more cost-effective than acquiring new ones. Therefore, reducing churn is vital for sustaining revenue streams, enhancing profitability, and maintaining market share.

Business Objectives

- To develop a predictive model that can accurately identify customers who are at risk of leaving the service in the near future
- Identify key drivers of churn
- Accurately predict at-risk customers
- Measure Impact of Retention Strategies



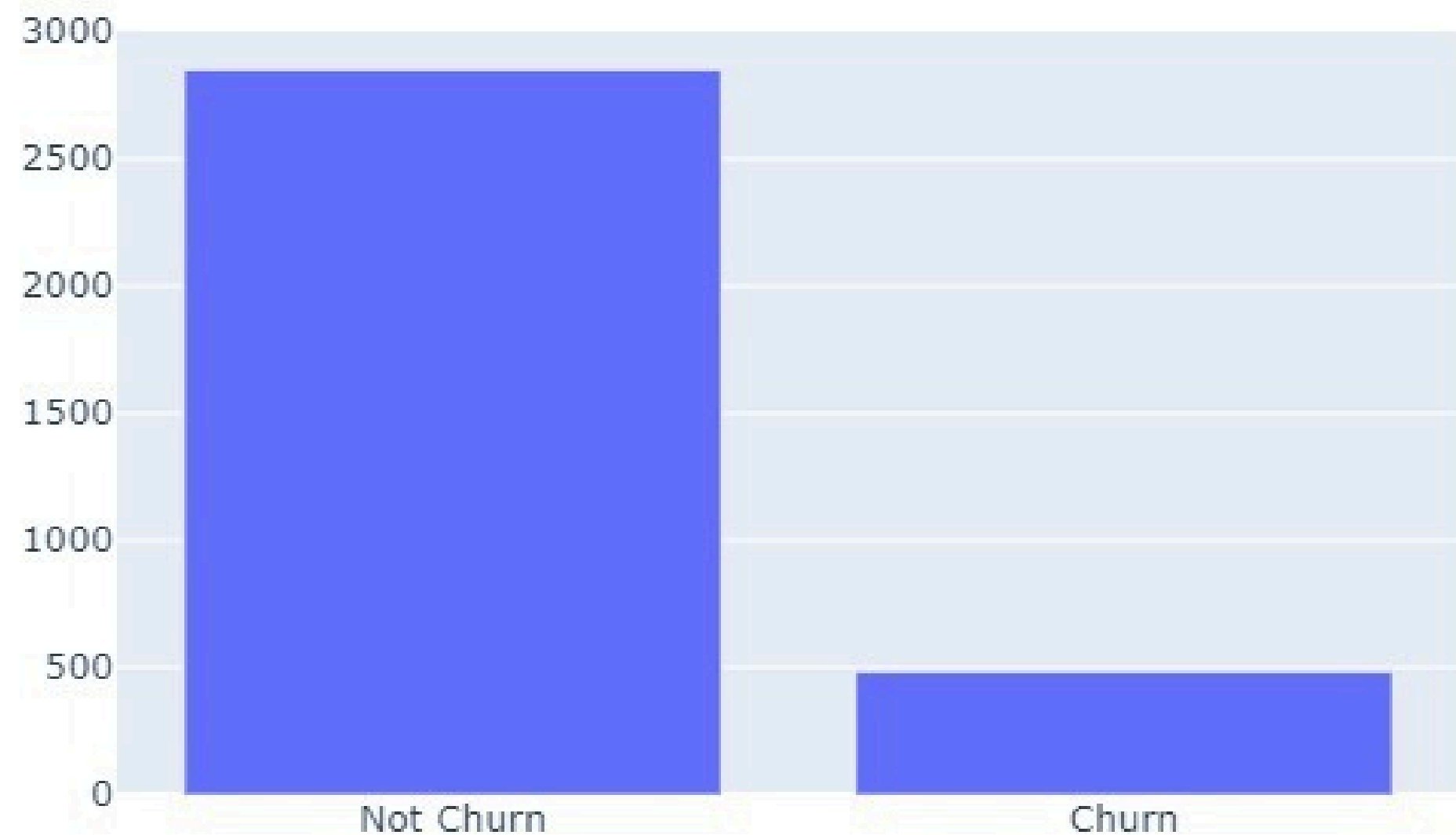
DATA UNDERSTANDING

- 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).



How many customers have churned?

Churn Distribution



Number of
customers :

3,333

Number of customers
that churned;

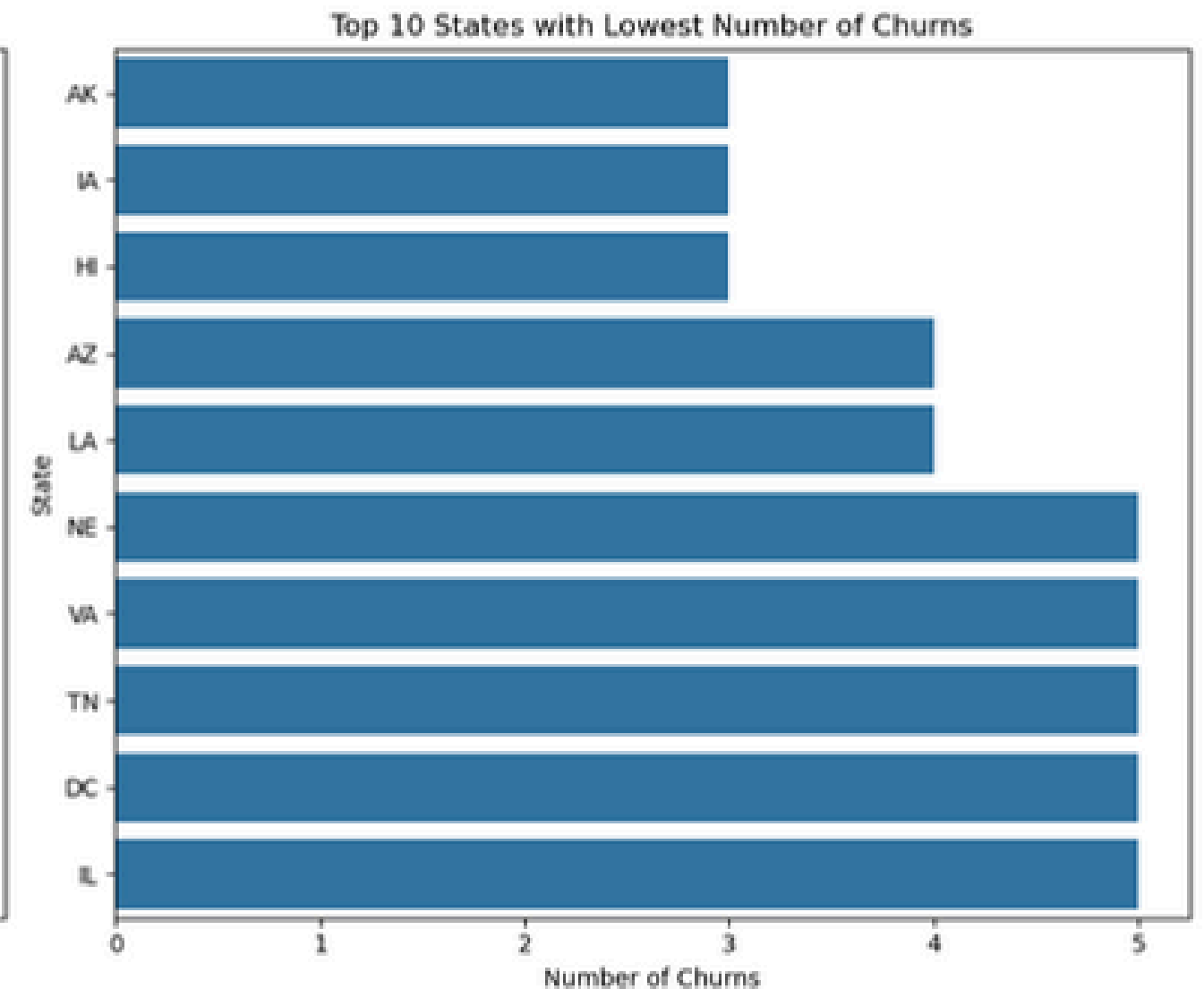
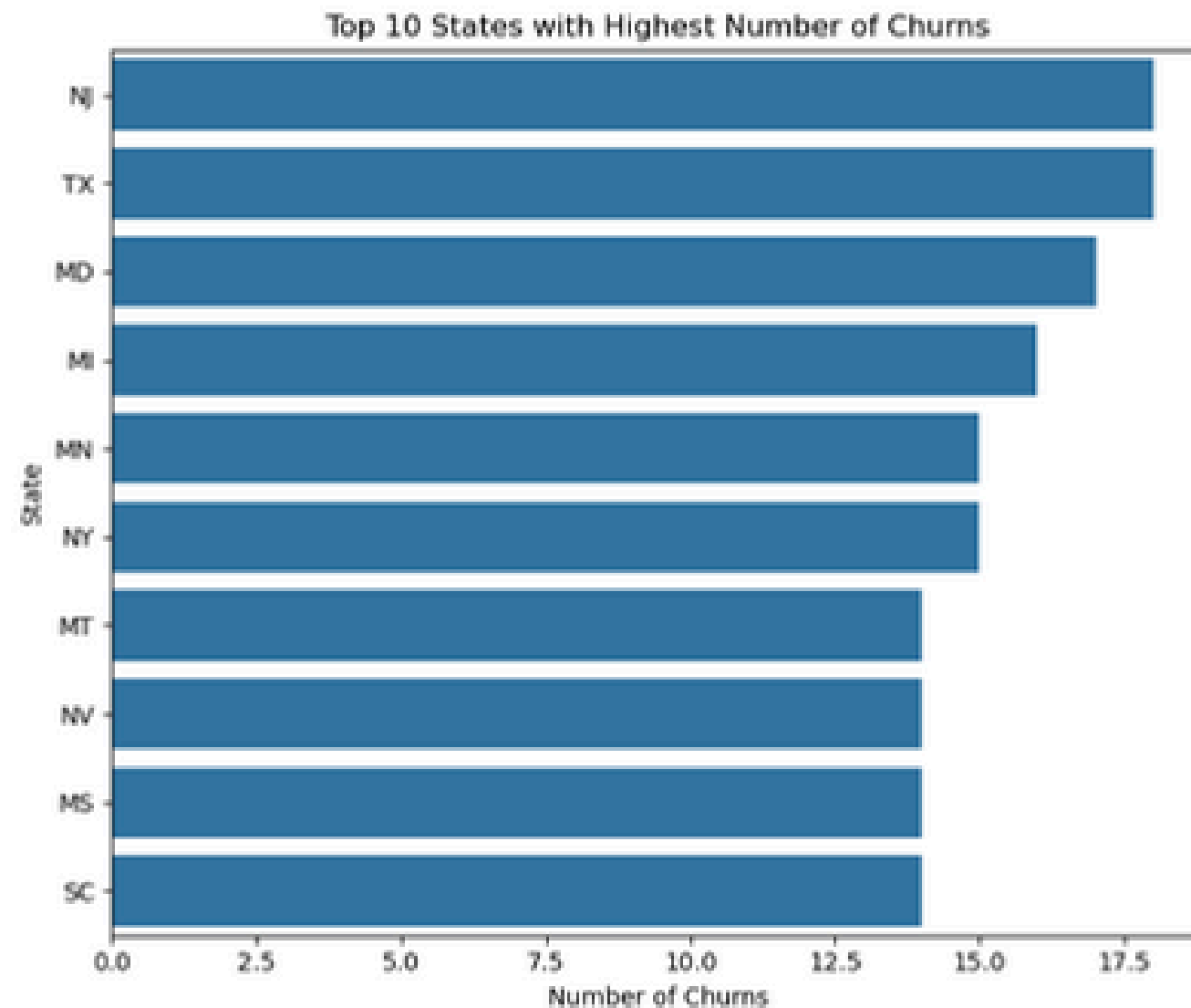
483

Churn Percentage;

14.5%

Which states have the highest and Lowest churn rate?

- New Jersey, Texas and Maryland have reported the highest customer churn numbers.
- Alaska, Iowa and Hawaii reported the lowest churn numbers.



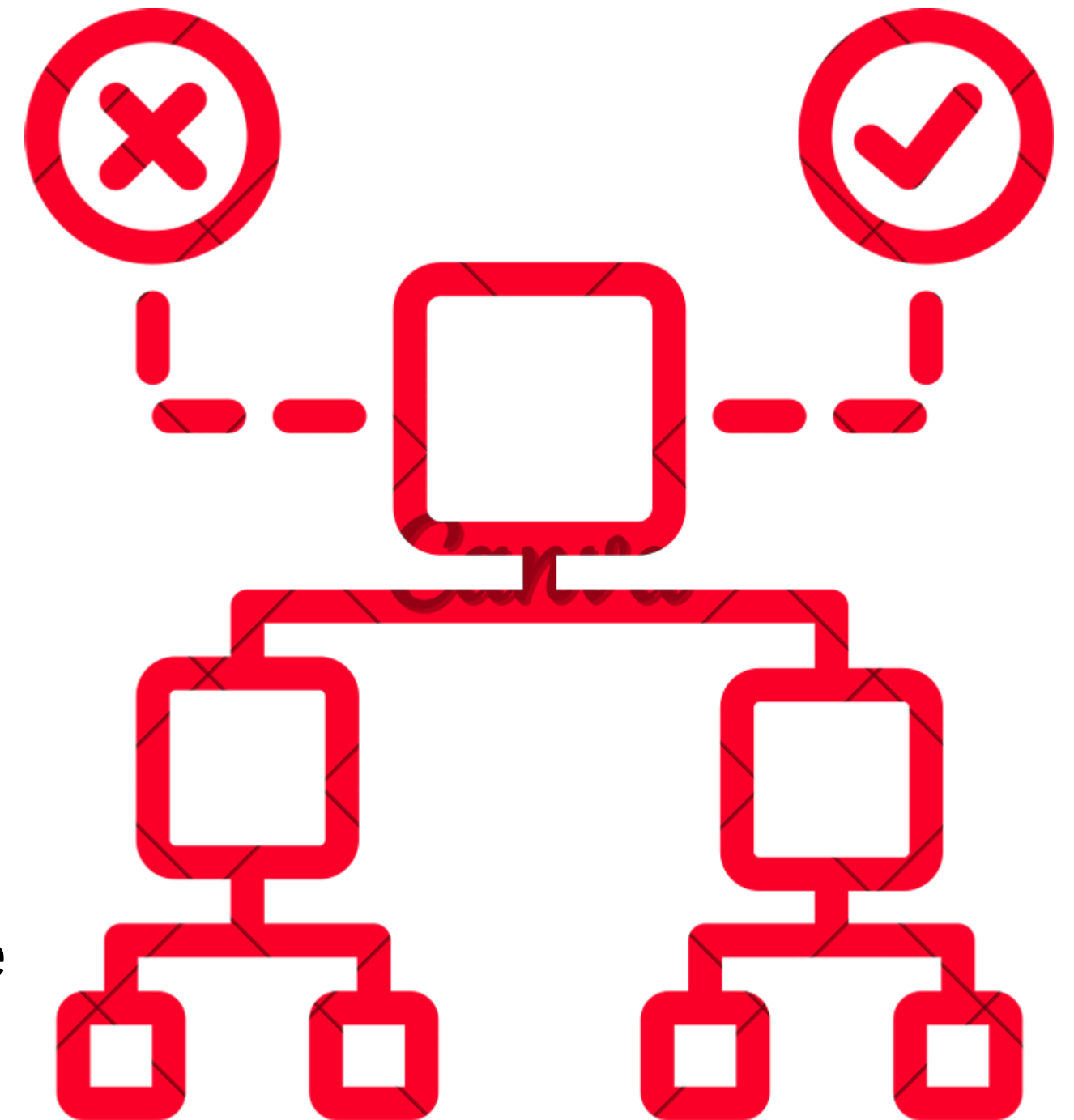
Modeling

Algorithms Used:

- Logistic Regression
- Decision Trees

Metrics Used:

- Recall - focusses on the model's ability to identify positive cases(churn)
- F1- score- focuses on balancing the false positives and false negatives to create a holistic score.
- AUC(Area Under the Curve) - provides a holistic measure of a model's ability to discriminate between the positive and negative classes, with values closer to 1 indicating a better performing model.



Evaluation



01

Recall Score of 69% : The model correctly identifies 69% of the customers who actually churned.

02

AUC (Area Under the Curve) is 0.79 : the model has a good ability to distinguish between customers who will churn and those who will not.



03

F1-score of 64% : the model strikes a balance between capturing true churners (recall) and ensuring that the identified churners are indeed likely to churn (precision). There may be a significant number of false positives or false negatives, impacting the overall effectiveness of the model.

Recommendations

01

Enhance Retention Strategies : Consider implementing targeted retention strategies for those not identified by the model. This might involve monitoring at-risk customers more closely or improving the model to capture a larger portion of potential churners.

02

Prioritize High-Risk Customers : Leverage the model to prioritize customers with a high likelihood of churn for retention efforts. Focus marketing and customer service resources on these high-risk customers to maximize the impact of retention initiatives.

03

Monitor Model Performance : Continuously monitor the model's performance over time. As customer behavior changes, the model may need adjustments to maintain its effectiveness. Regularly track key metrics like recall, AUC, and F1-score, and be prepared to retrain or recalibrate the model as needed.

Thank
you!