



Syriatel telecommunica tion

Customer churn prediction



Table of contents

- Introduction
- Problem statement
- Objectives
- Patterns and findings
- Conclusions
- Recommendations



Introduction

Customer churn refers to the loss of customers or subscribers for any reason at all.

- The telecommunications industry is competitive and characterized by very high churn rates.
- Attracting new customers is both difficult and expensive hence avoiding churn, should be given a higher priority than trying to attract new customers.





Problem Statement

- Syriatel has a significant problem with customer churn.
- To solve this problem, it needs to identify the customers who are likely to churn and try to keep them.
- Machine learning models can help telco's in predicting the customers who are most likely to churn, based on various factors such as customer usage patterns, payment history, and demographics.
- The goal is to build a classifier to predict whether a customer will ("soon") stop doing business with Syriatel so that the company can take action to retain them.




Objectives

- Build a model that predicts customer churn with high recall.
- Identify the features that influence customer churn.
- Provide recommendations to reduce customer churn.

Data used

- The data set was obtained from Kaggle.
- It contained features such as customer demographics, their usage patterns and subscriptions and payments which can be used to predict customers who are most likely to churn.
- The dataset comprises 21 columns with a total of 3333 entries.

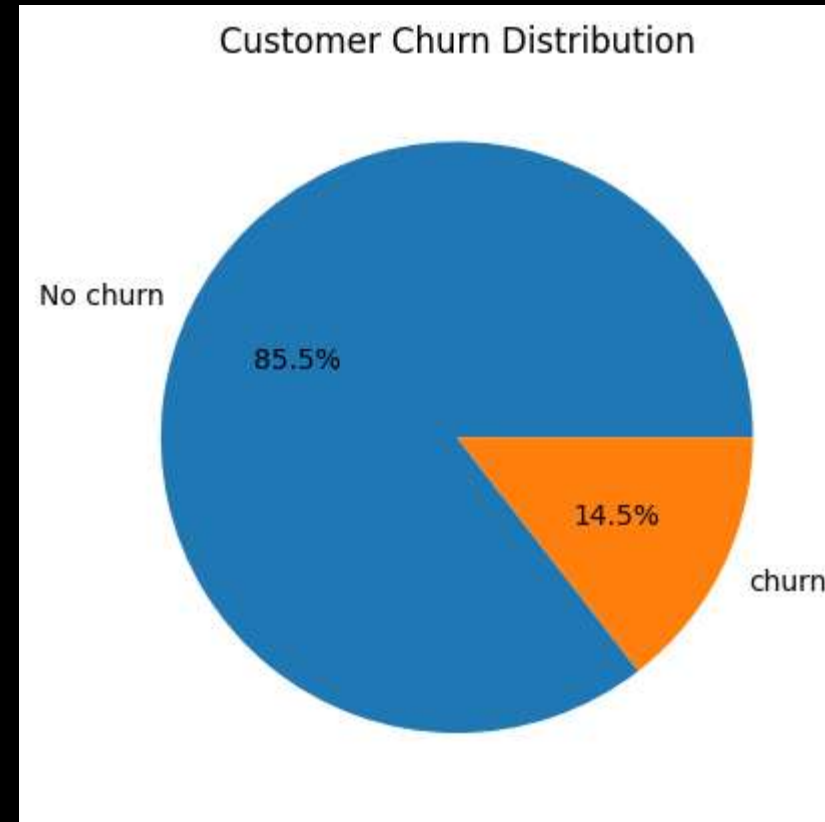


The background is a dark blue gradient. In the lower-left and center, there are glowing blue lines and nodes, resembling a network or data flow. Overlaid on this are several data visualization elements: a bar chart with blue bars of varying heights, a line graph with a blue line showing an upward trend, and a faint city skyline in the background. The text 'Patterns and findings' is written in a white, monospaced font on the right side.

Patterns
and
findings

Customer churn distribution

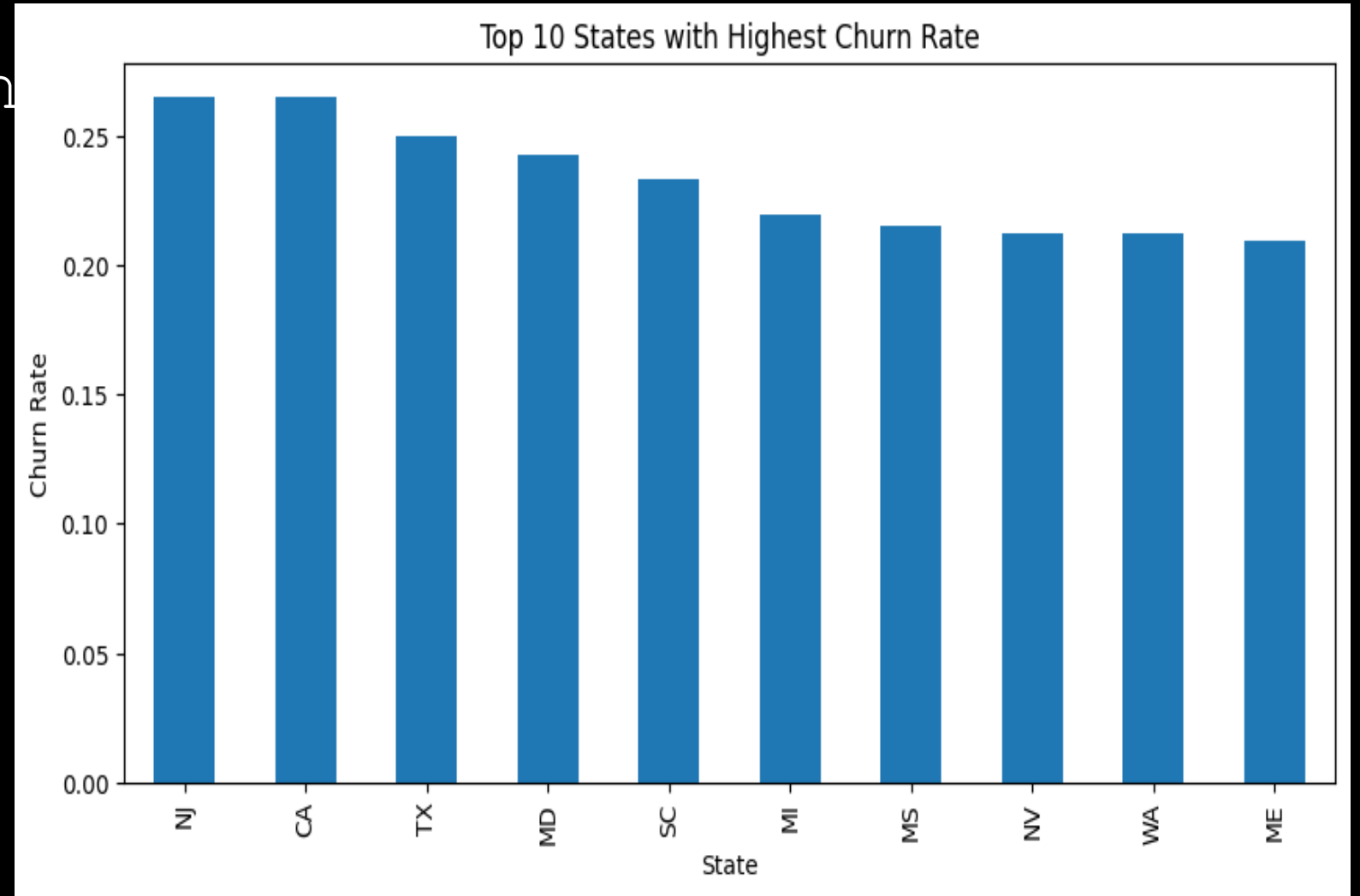
From the chart: 85.5% of customers did not churn (represented by the blue section), meaning they continued to use the product or service. 14.5% of customers did churn (represented by the orange section), meaning they stopped using the product or service



State and churn rates

The top 5 states with the highest churn rate are:

- NJ (New Jersey)
- CA (California)
- TX (Texas)
- MD (Maryland)
- SC (South Carolina)



Modelling

The data was trained using the following machine learning models:

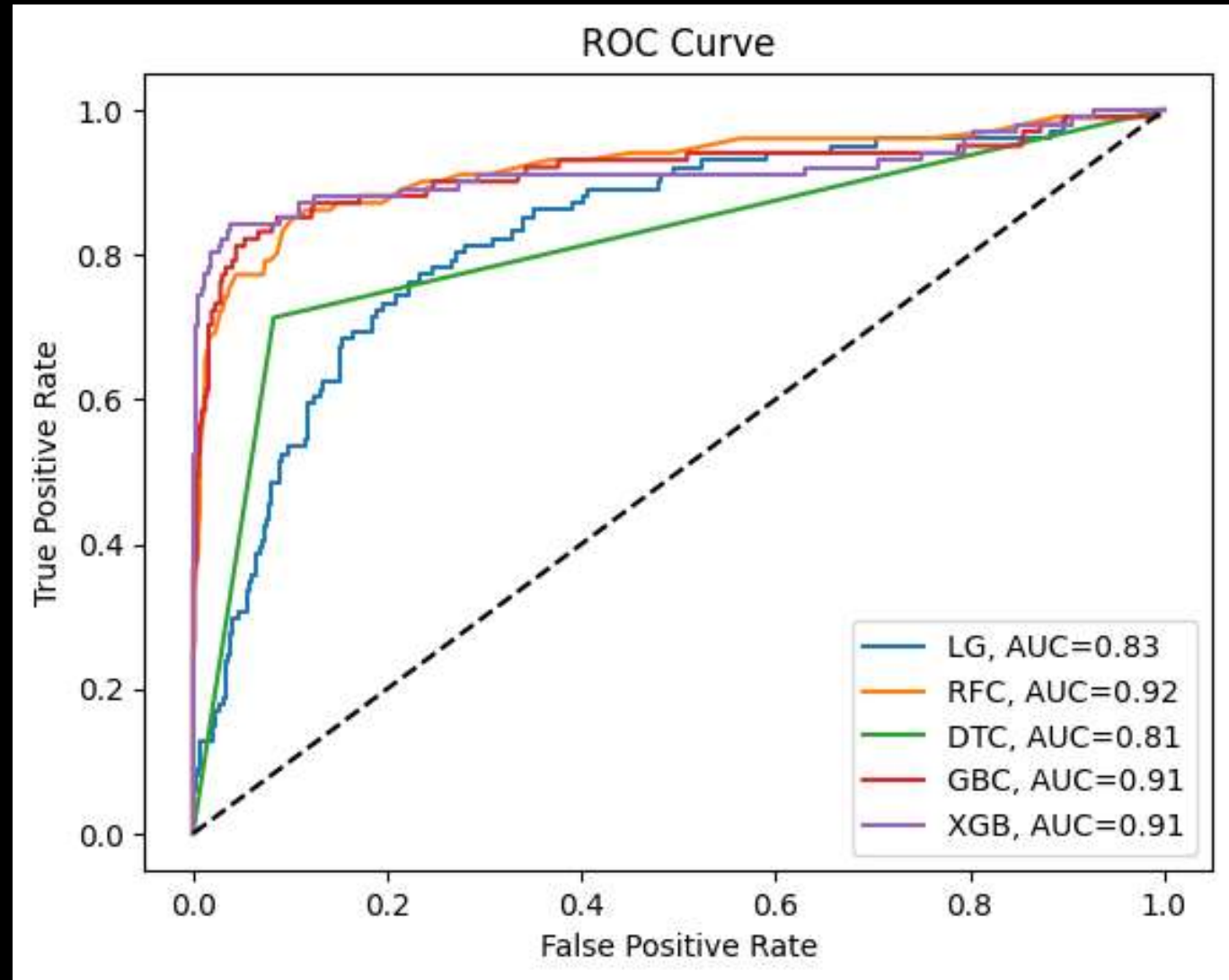
- Logistic Regression.
- Random Forest Classifier.
- Decision Tree Classifier.
- XGB Classifier.



Model evaluation

- Logistic Regression (LG): 0.83
- Random Forest Classifier (RFC): 0.92
- Decision Tree Classifier (DTC): 0.81
- Gradient Boosting Classifier (GBC): 0.91
- XGBoost Classifier (XGB): 0.91

Random Forest Classifier (RFC) with an AUC of 0.92 is the best model to predict customer churn among the compared models



A close-up, blurred image of a pen writing on a document. The document features a line graph with a dotted trend line. The pen is positioned at the top right, and the graph shows a series of peaks and valleys, with the dotted line indicating an overall upward trend.

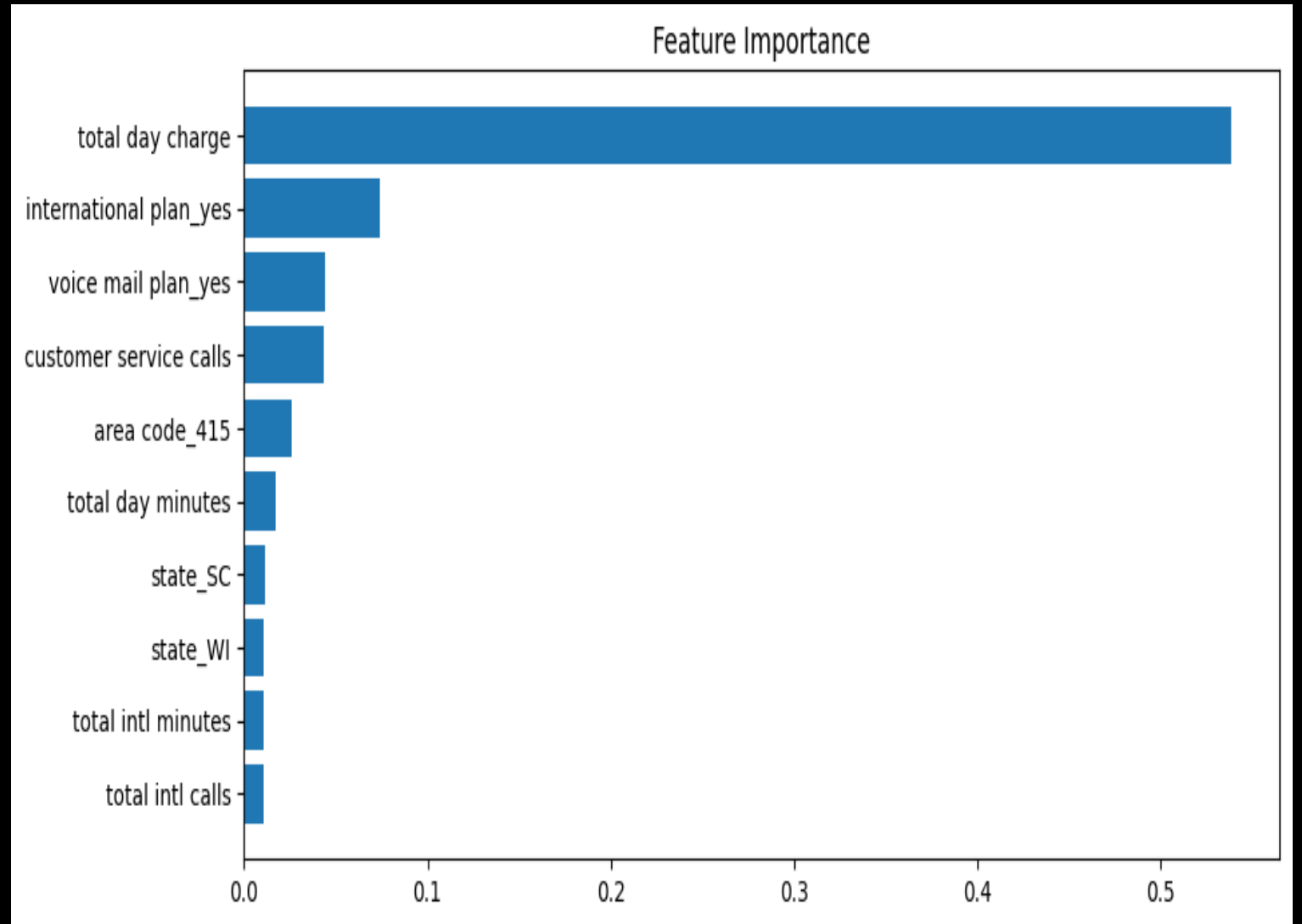
Model Evaluation

- The XGBoost model achieved a recall of 0.79 for churned customers. This means it correctly identified 79% of the actual churn cases.
- The XGBoost Classifier ROC score of 0.91 indicates a high level of accuracy in predicting customer churn. This means that the model is able to correctly identify 91% of the cases where customers are likely to leave the service

Top five features

The top 5 features that affect customer churn include:

- Total Day Charge
- International Plan
- Voice Mail Plan
- Area Code (415)
- Customer Service Calls



A woman with dark skin and short hair, wearing a yellow sweater, is shown in profile, looking upwards and to the left. She is interacting with a futuristic, glowing blue digital interface. The interface features a large globe on the left, various data charts, graphs, and icons floating around her. The background is dark, making the glowing elements stand out.

Conclusion

- By leveraging predictive analytics like customer churn prediction, we can proactively address potential issues and tailor the company's services to meet individual customer needs, thereby enhancing customer satisfaction and loyalty.

Recommendations

- XGBoost for predicting customer churn
- Enhance international service offerings and review pricing.
- Improve voice mail services and offer tutorials or assistance.
- Customer care team should strengthen customer support, resolve issues effectively, and enhance service quality.
- Improve network quality

