SyriaTel Telecom's Customer Churn Prediction

Daniel Akwabi

OVERVIEW

Customer churn (the rate at which customers stop doing business with a company) directly impacts revenue, profitability and long term growth for telecommunications companies. A spike in churn rates would affect the overall performance of SyriaTel Telecom.

By utilizing existing customer information, I will do a forecast on probable exits by customers. In turn, SyriaTel will be in a position to optimize customer retention strategies, maintain customer loyalty and sustain competitive advantage by establishing the churn trends

BUSINESS PROBLEM

01

SyriaTel is experiencing customer churn leading to lost revenue, increased customer acquisition expenses and a reduction in market share.

02

SyriaTel is unable to formulate retention plans due to lack of a clear understanding of customer churn and high-risk customers prediction leading to missed opportunities to improve customer experience, optimize retention and preserve long term profitability.

03

This calls for a research on customers who are likely to exit SyriaTel.

OBJECTIVES

01

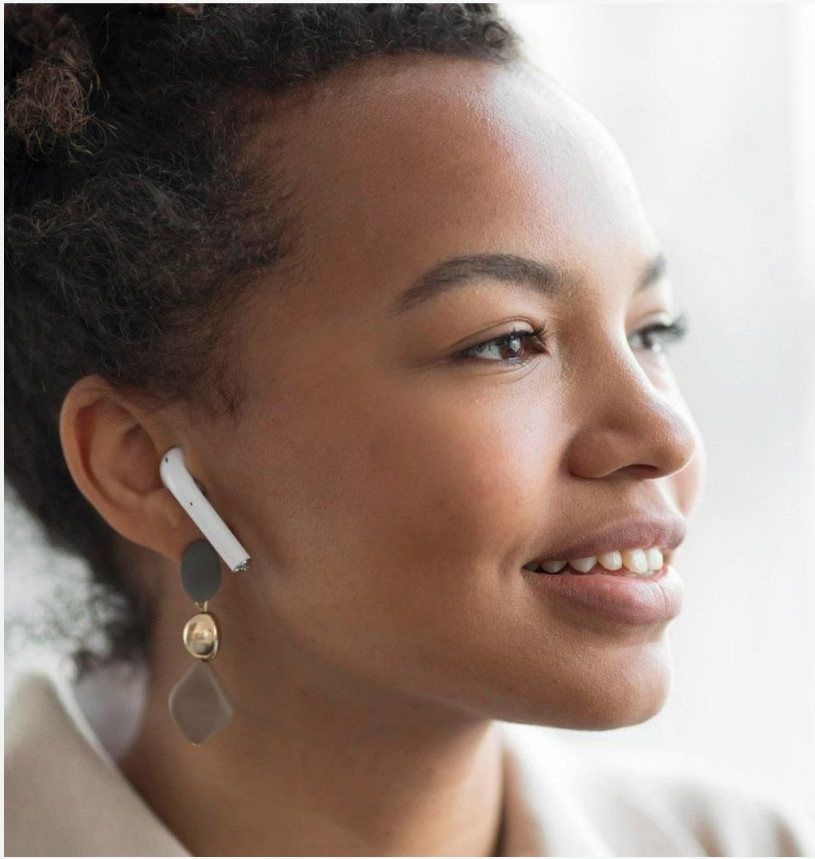
Identify high risk customers before they churn

02

Highlight key factors driving attrition

03

Lay out actionable insights to facilitate SyriaTel's implementation of selected retention strategies.

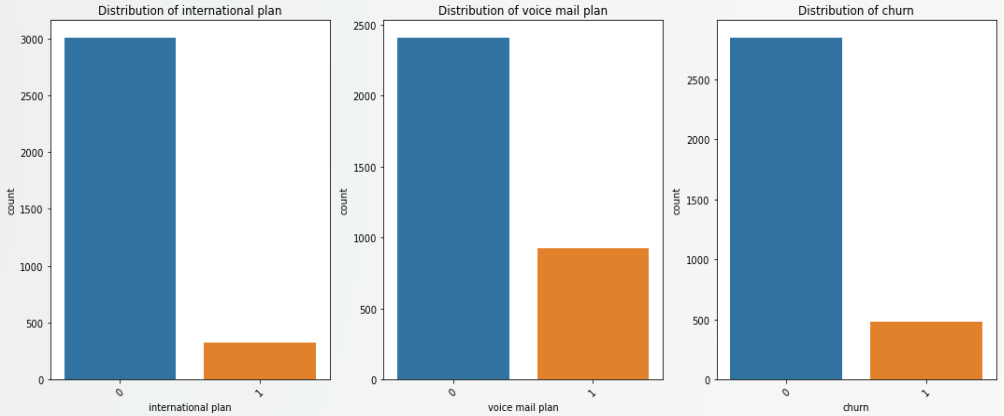


DATA UNDERSTANDING

The dataset includes details on customer behavior, usage metrics, and subscription plans, with the goal of predicting **customer churn**.

Churn is our targeted variable in this instance.

Additional features cover account tenure, international plan status, voicemail plan status, voicemail count, total usage minutes, call summaries as well as customer service interactions.



MODELLING

Two models were used in this context :

1. Logistic regression model
2. Random Forest classifier

EVALUATION

1. Logistic regression model

●**Accuracy**: The model accurately predicted 75% of the cases overall. Although, accuracy can be misleading hence it is not the best metric to evaluate performance in imbalanced datasets.

●**ROC-AUC score = 0.7891** The score is good and indicates that the model has a solid ability to distinguish between the two classes. Its closeness to 1 (ideal) implies the model performs well at ranking predictions.

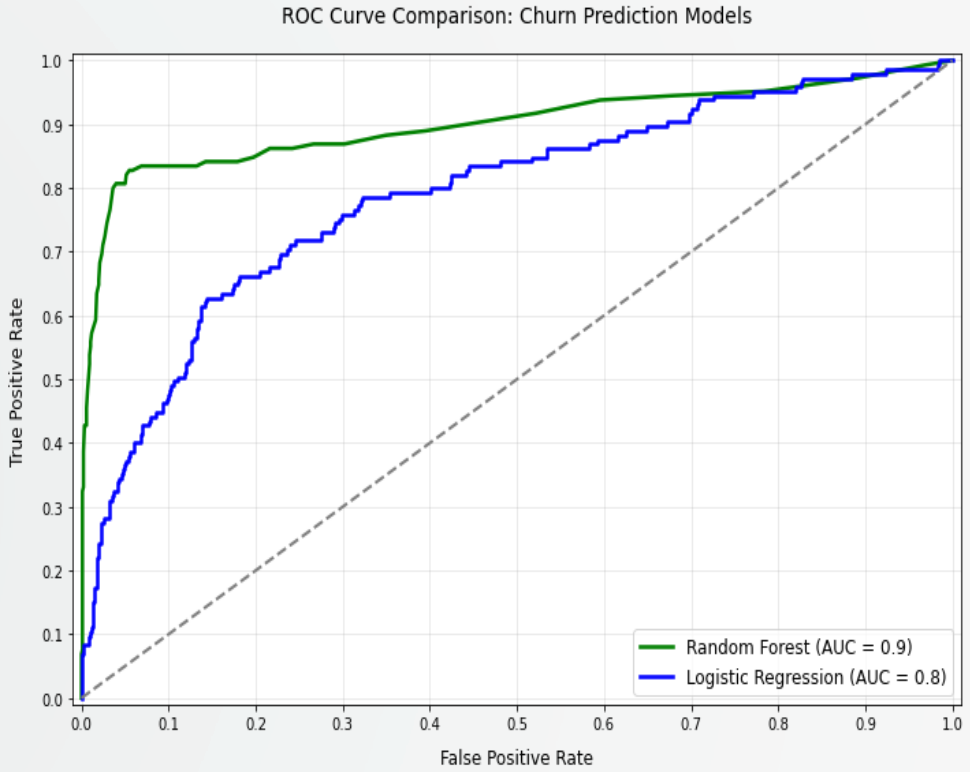
2. Random Forest Classifier

- **Accuracy (0.92)**: The model precisely predicts the probability of a customer churning or not in 92% of cases.
- **ROC-AUC score(0.9002)** It elaborates excellent model performance due to its high capability of distinguishing between churn customers and non-churn customers.

COMPARISON BETWEEN LOGISTIC REGRESSION & RANDOMFOREST CLASSIFIER

Evaluation

The more skewed the ROC curve is to the top-left corner, the better the model's performance. In this case, the **Random Forest** curve is consistently above the Logistic Regression curve.



FEATURE IMPORTANCE ANALYSIS

- **Most Important Features:**

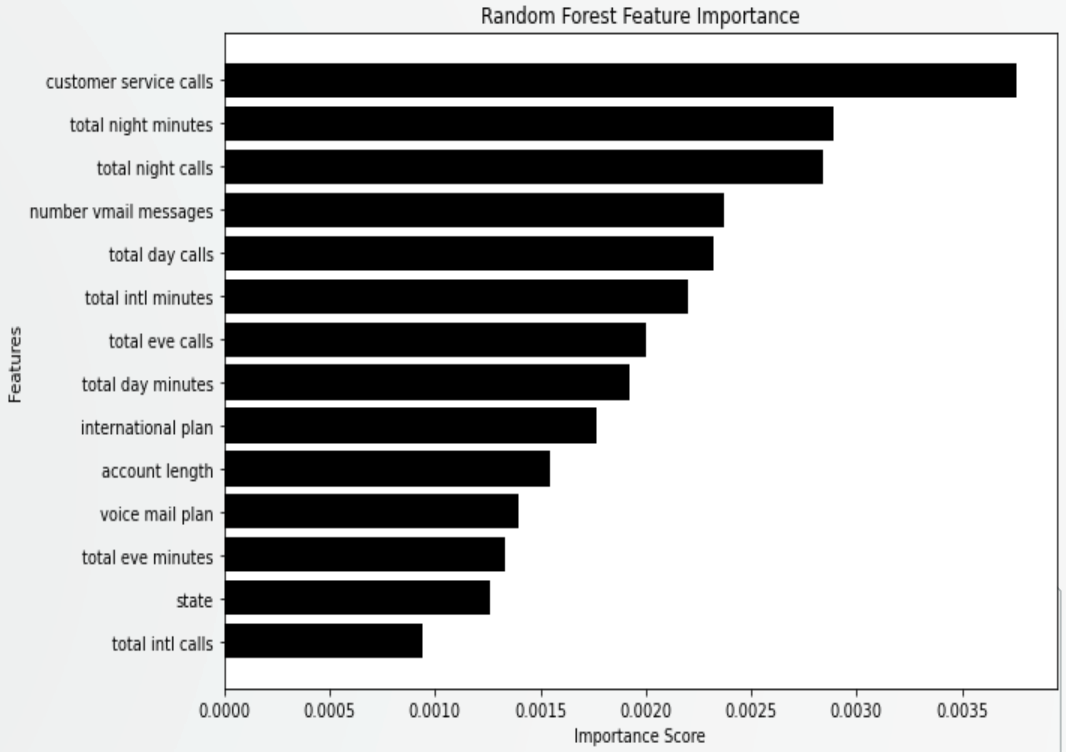
Customer service calls and total night minutes are the top two most important features, with the highest importance scores.

- **Moderately Important Features:**

Features like total night calls, number of voicemail messages, and total day calls have intermediate importance,

- **Least Important Features:**

Total eve minutes, state and total intl calls are at the bottom of the list,



CONCLUSIONS

Logistic Regression is precise on non-churn predictions but not for churn prediction as highlighted by the low precision and recall for churn.

Random Forest, provides a balanced performance across both classes, as it is better at handling of class imbalance.

BUSINESS RECOMMENDATIONS

- **For high-value customer retention:**

Use Random Forest to identify at-risk customers.

- **For regulatory/compliance scenarios:**

Consider Logistic Regression for its interpretability.

- **Optimal threshold selection:**

Balance between false positives and false negatives

Use precision-recall curves for imbalanced data

NEXT STEPS

- Advance the model with updated data
- Utilize the Random Forest model in their operations
- Implement customer retention strategies courtesy of the prediction models

THANK YOU!

Email me on **danakwabi@gmail.com**

Check out my GitHub **@daniel-akwabi**