CUSTOMER CHURN PREDICTION FOR SYRIATEL COMPANY

January 2025

otienogloryann790@gmail.com

www.syriatelecommmunications.com

CONTENT

OVERVIEW

The problem identified is customers are leaving (churning), and the company wants to predict and reduce this.The main goal is to build a model to identify customers at risk of churning so the company can take action. Business impact will be reducing churn to improve customer retention and aid the company in reducing how much money was lost because of customers who churned.

BUSINESS AND DATA UNDERSTANDING

The dataset consists of customer information, including details on call charges, service usage, account length, and customer service interactions, along with their churn status.

Key features such as total day charge, customer service calls, and international plan usage appear to have a significant influence on whether a customer will churn. Initial observations suggest that customers with high service calls or higher charges are more likely to churn, indicating potential dissatisfaction with pricing or service quality.

Understanding these patterns helps in selecting relevant features for modeling and improving the accuracy of churn predictions.

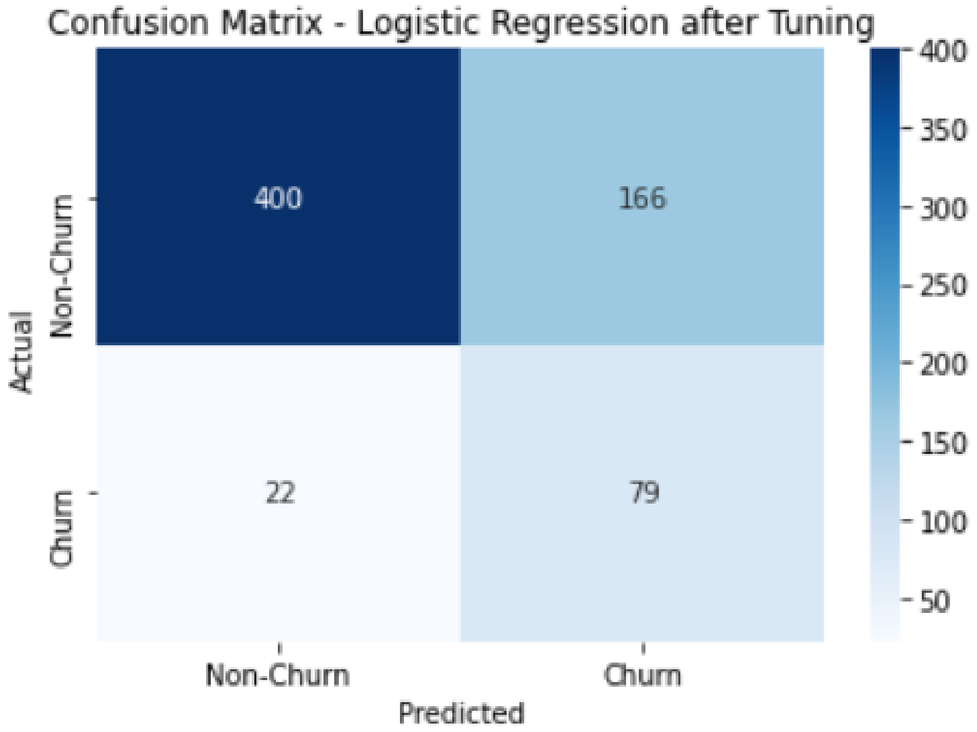
MODELING

The goal was to find the most accurate model for predicting churn.

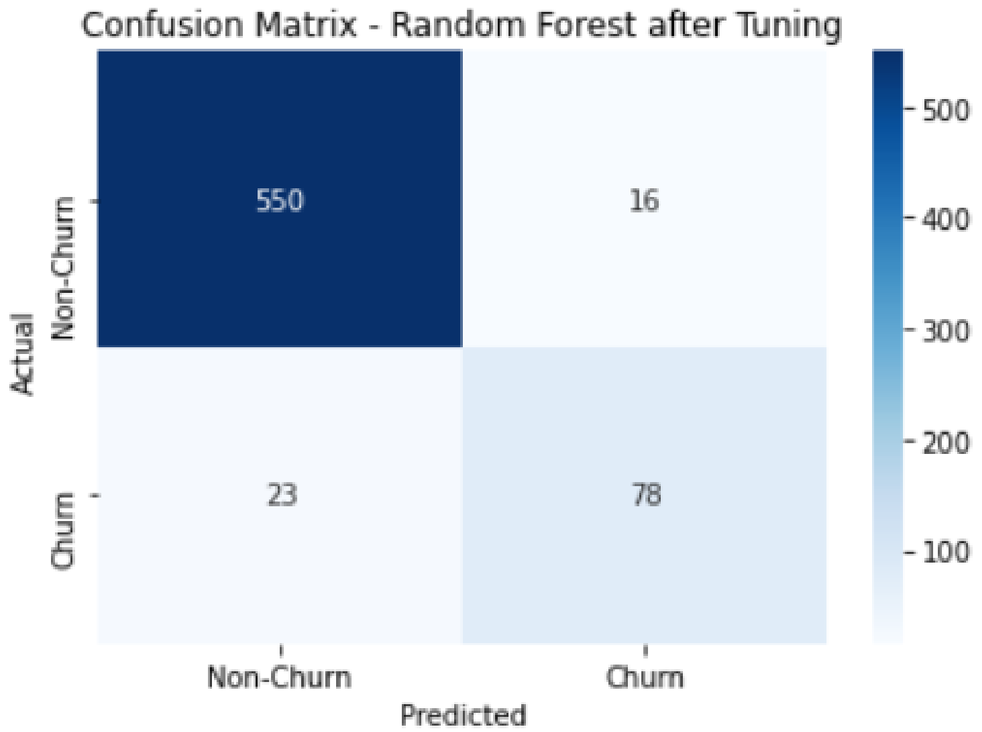
- **Models Tested :** Logistic Regression and Random Forest.
- **Best Model Chosen :** Random Forest performed best due to its ability to capture complex patterns in the data.
- **The Random Forest model** is the best choice because it predicts customer churn more accurately and helps us understand which factors influence customers to leave. This will allow us to take better actions to reduce churn.

Model	Accuracy	Recall Score	F1 Score
Logistic Regression	0.718141	0.782178	0.456647
Random Forest	0.941529	0.772277	0.799999

EVALUATION

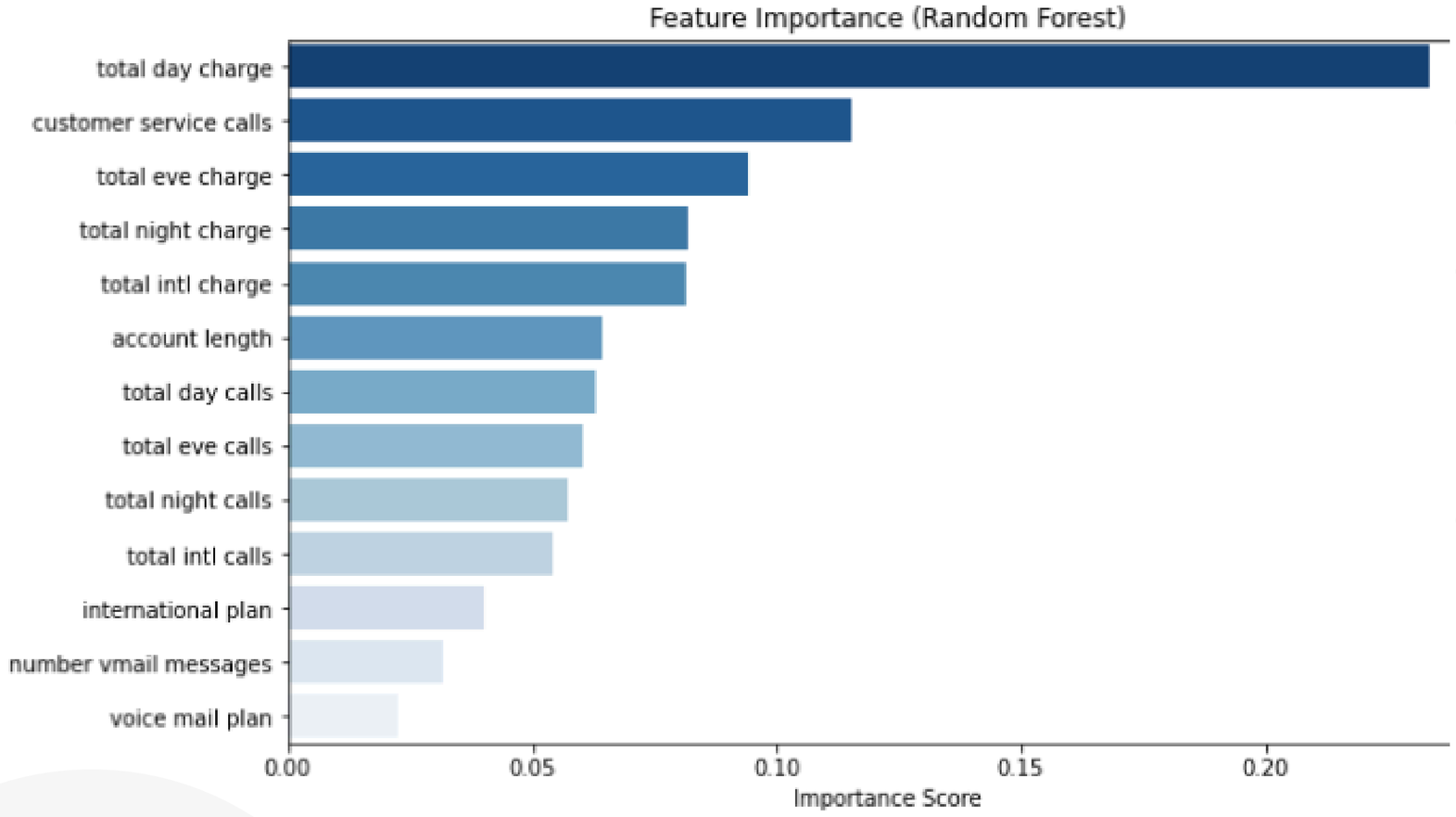


The model achieved slightly higher recall (78.2%), meaning it is better at catching actual churners. This makes it a strong choice if the primary goal is to identify and retain as many churners as possible. However, it comes at the cost of high false positives (166 non-churners misclassified), leading to potentially wasted retention efforts.



The model had a slightly lower recall (77.2%), meaning it missed one extra churner compared to Logistic Regression. However, it significantly reduced false positives (only 16 non-churners misclassified). This means less unnecessary intervention and better use of retention resources.

FEATURE IMPORTANCE



- **Total Day Charge** is the most significant factor influencing churn.
- **Customer Service Calls** is also highly correlated with churn, meaning customers who call support frequently are more likely to leave.
- Other charges **(evening, night, and international)** and account-related factors (**account length, total calls)** also contribute but with lower impact.

RECOMMENDATIONS

- **Personalized Retention Offers:** Identify high-risk customers and offer tailored incentives like discounts or exclusive deals to encourage them to stay.

- **Proactive Customer Support**: Since many churned customers had multiple customer service calls, improve support by resolving issues faster or introducing a loyalty support team.

- **Service Plan Optimization**: Offer flexible plans based on customer usage patterns to reduce dissatisfaction, especially for high-usage customers with high day charges.

NEXT STEPS

- **Implement a Churn Prevention Strategy**: Use the insights from feature importance (e.g., high day charge, customer service calls) to develop targeted retention campaigns.

- **Continuously Monitor and Improve the Model**: Regularly update the model with new data to maintain accuracy and adapt to changing customer behavior.

THANK YOU

FOR YOUR ATTENTION

January 2025

otienogloryann790@gmail.com

www.syriatelecommmunications.com