USING MACHINE LEARNING TO PREDICT CUSTOMER CHURN AT SYRIATEL MOBILE TELECOM

PREPARED BY LYNETTE WANGARI

OVERVIEW

- In the competitive telecommunications market, Syriatel faces significant challenges in retaining its customer base and ensuring high satisfaction levels. The company struggles with identifying at-risk customers and addressing their issues proactively, which is crucial to maintaining a competitive edge. Without effective data analytics and machine learning, Syriatel risks losing market share to competitors who better leverage these technologies to enhance customer experiences and loyalty.
- This project aims to address customer churn and improve satisfaction for Syriatel. By implementing advanced data analytics and machine learning, Syriatel can develop strategies to identify at-risk customers, understand their reasons for potential departure, and deploy targeted interventions. This approach will help reduce churn, enhance customer satisfaction, and ensure long-term success and competitiveness in the telecommunications market.

BUSINESS UNDERSTANDING

- Syriatel Mobile Telecom, facing competitive challenges, emphasizes the need to enhance customer satisfaction and retain its subscribers. To mitigate the potential threat of customer churn, the company aims to use data analytics to identify at-risk customers and tailor strategies to reduce churn risk. This approach is crucial for cutting customer turnover costs and fostering long-term loyalty, driving sustainable business growth.
- The proposed model offers significant benefits for all stakeholders. By reducing customer churn rates, the company can increase revenues, profits, and market sustainability. Customers will enjoy improved services and support, while shareholders can expect higher returns on investments. Additionally, employees may benefit from better remuneration packages and bonuses as the company prospers, ensuring mutual growth and prosperity for all involved.

RESEARCH QUESTIONS AND OBJECTIVES

Research Objectives:

- To identify the key features that determine if a customer is likely to churn.
- To determine the most suitable model to predict Customer Churn.
- To establish Customer retention strategy to reduce churn

Research Questions:

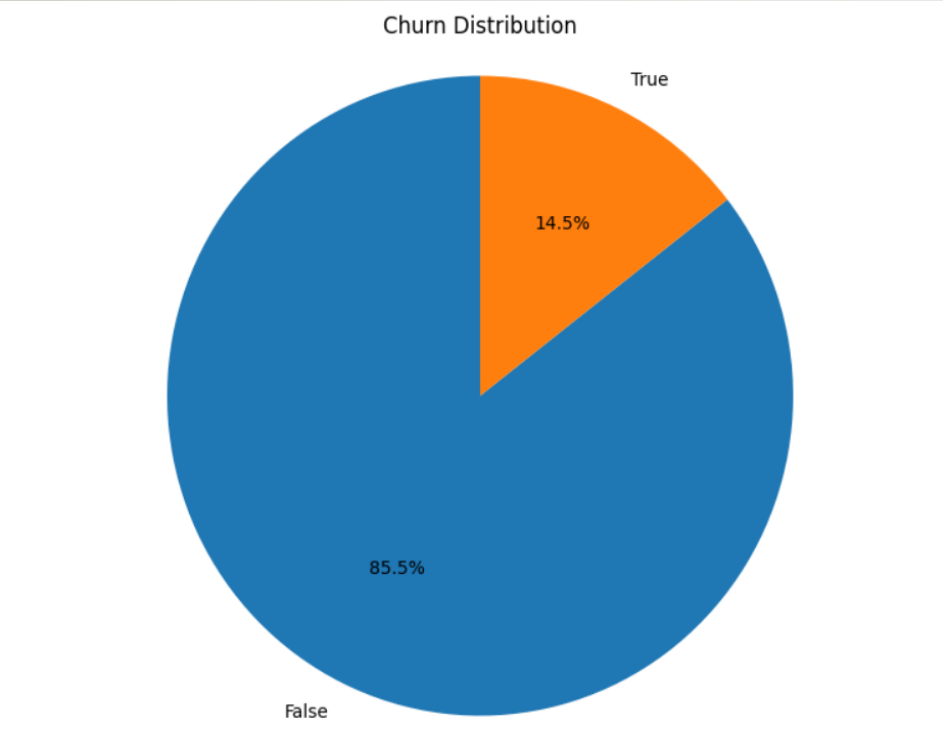
- What are the most significant predictors of customer churn for Syriatel Mobile Telecom?
- Which Machine Learning Model is the most suitable in predicting Customer Churn?
- What strategies can Syriatel Mobile Telecom implement to retain customers and reduce churn rates?

DATA UNDERSTANDING

- The churn in Syriatel dataset contains information about customer activity and whether or not they canceled their subscription with the firm.
- The dataset contains 3,333 entries and 21 columns, including information about the state, account length, area code, phone number, international plan, voice mail plan, number of voice mail messages, total day minutes, total day calls, total day charge, total evening minutes, total evening calls, total evening charge, total night minutes, total night calls, total night charge, total international minutes, total international calls, total international charge, customer service calls and churn.
- Each row having a the telephone number as the unique identifier.

EXPLORATORY DATA ANALYSIS

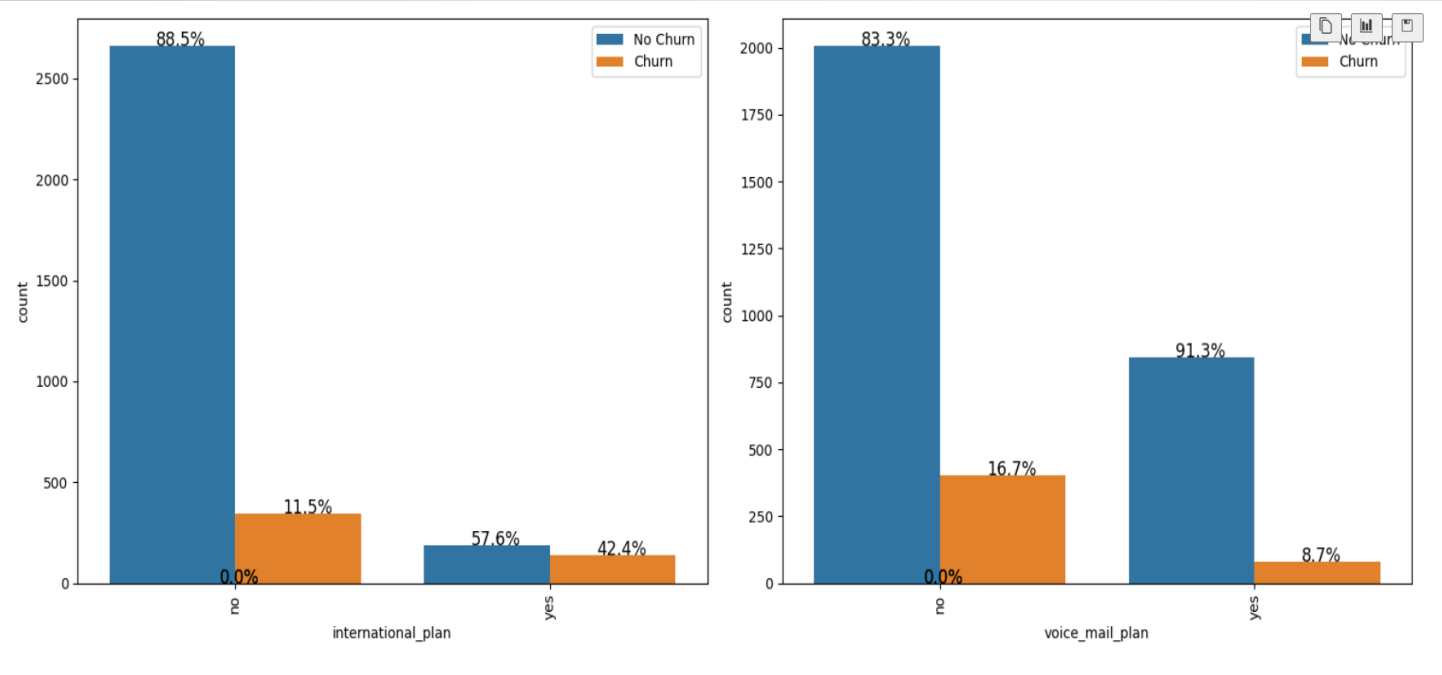
UNIVARIATE ANALYSIS



- Of the 3,333 customers in the dataset, 483 have terminated their contract with Syriatel. That is 14.5% of customers lost.
- True – Churned
- False – Not Curned

EXPLORATORY DATA ANALYSIS CONT'D

BIVARIATE ANALYSIS

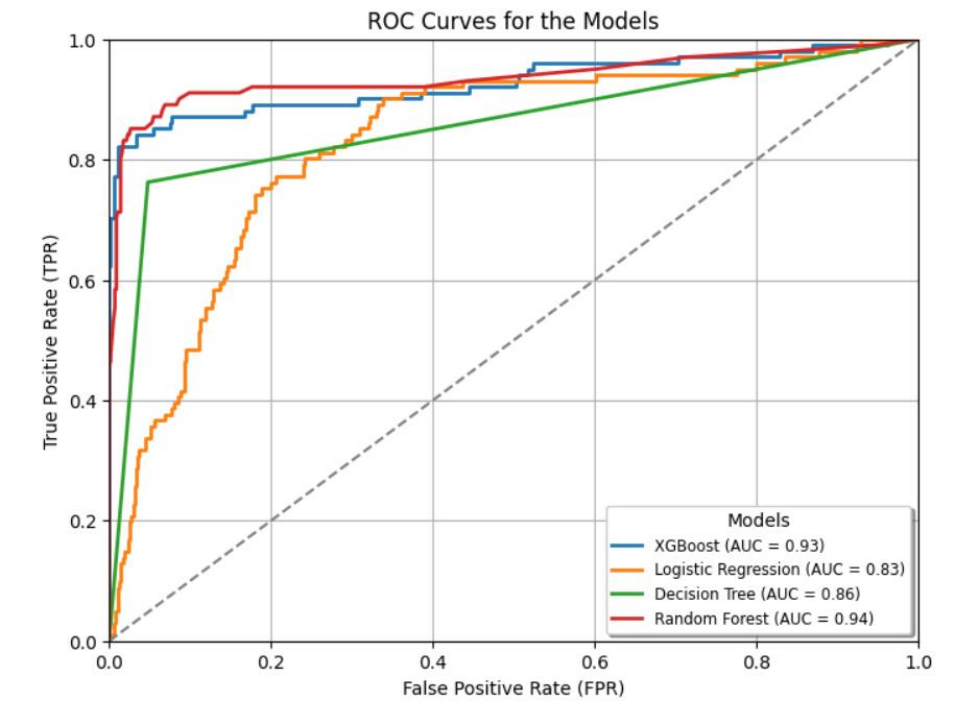


- For the international plan, a higher percentage of the customers who subscribed to the plan churned at 42.4% compared to those who did not subscribe to the plan at 11.5%. This suggests that there's a likelihood of churning after subscribing to the international plan.
- For the voice mail plan, a lower percentage of customers who subscribed to the plan churned at 8.7% compared to those who did not subscribe to the plan at 16.7%. This suggests that subscribing to the voice mail plan may be associated with a lower likelihood of churning.

MODELLING

- Four models were used to predict churn and further improvements were made to achieve the best predictive results. The models used were:
    - Logistic Regression
    - Decision Trees
    - Random Forests
    - XGBoost

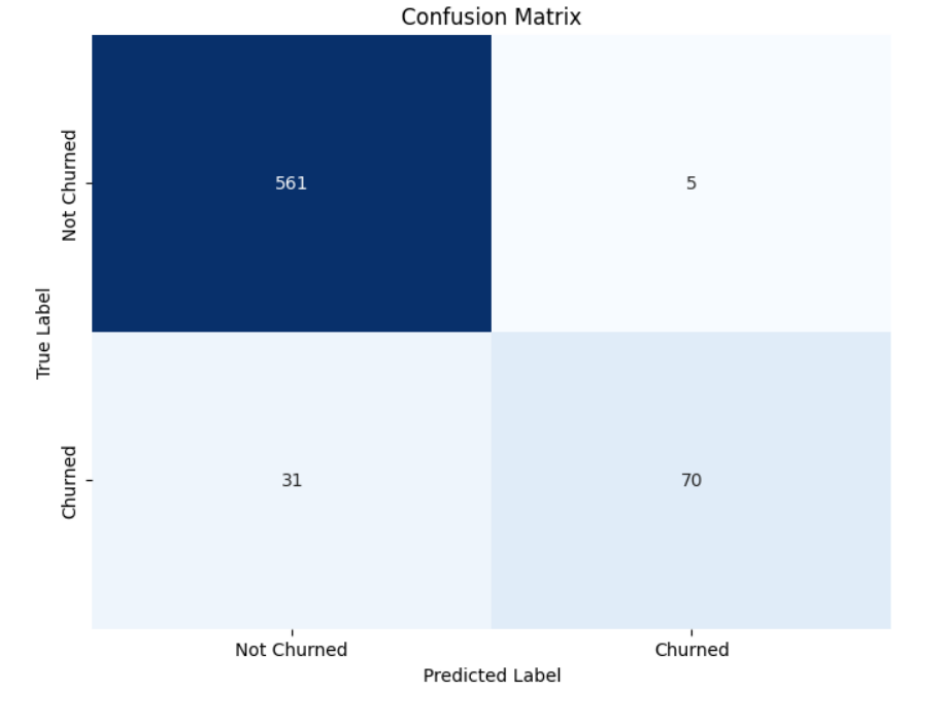
MODEL PERFORMANCE



The Random Forest model outperformed the others, showing a higher Area Under the Curve (AUC) of 0.94 and better classification performance, making it the most effective model for the given task.

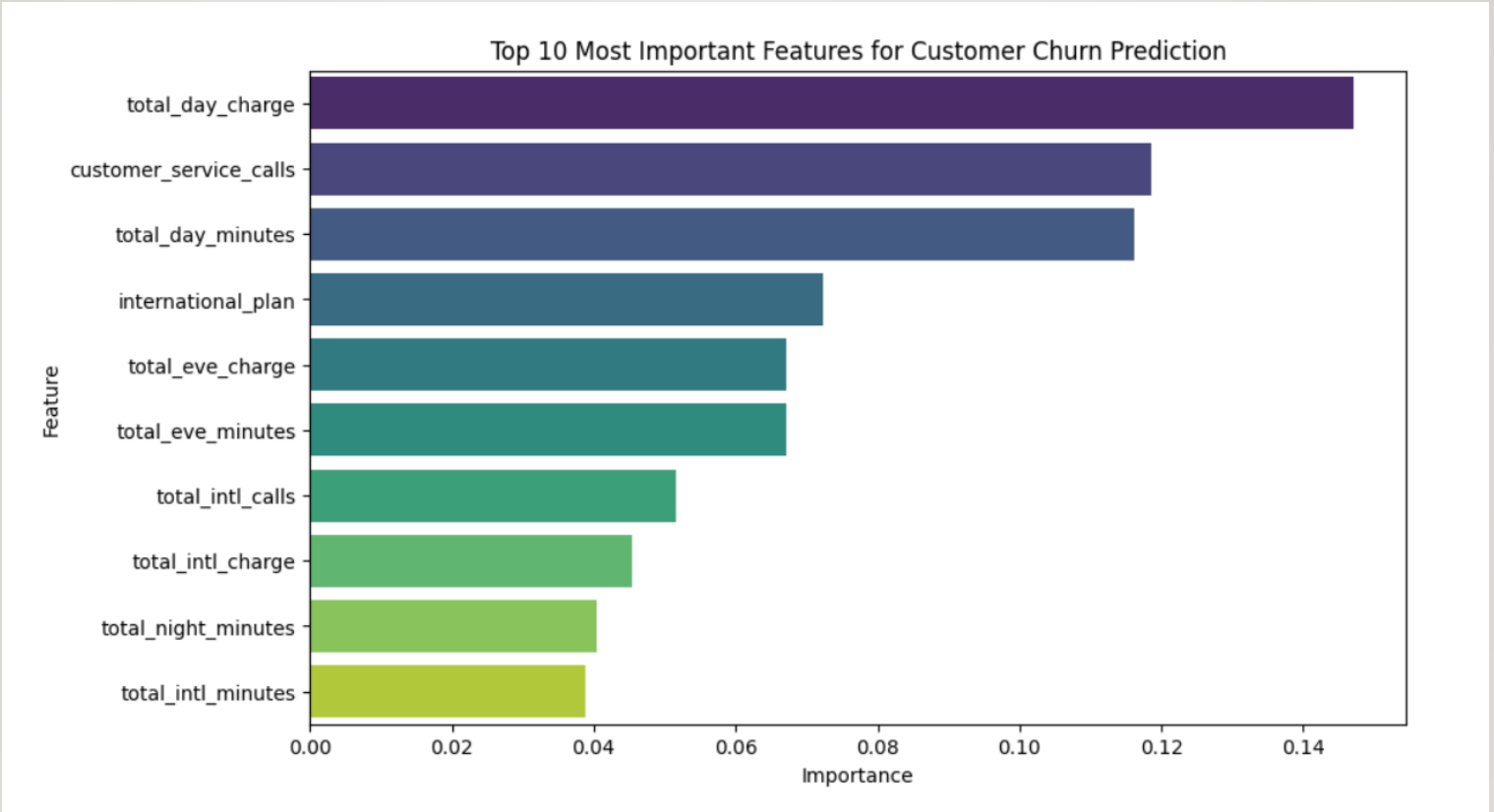
The XGBoost model also performed well with an AUC of 0.93, while Logistic Regression and Decision Tree models had lower AUCs of 0.83 and 0.86, respectively.

RANDOM FOREST MODEL METRICS



- **Cross-Validation Accuracy:** 95.05%
- **Precision:** 95.77%
- **Recall:** 67.33%
- **F1 Score:** 79.07%
- **Train Score:** 97.86%
- **Test Score:** 94.60%
- **AUC:** 0.94

FEATURE IMPORTANCE



CONCLUSIONS

- The analysis concludes that customer churn can be accurately predicted using machine learning models, with the Random Forest Classifier emerging as the recommended model due to its superior overall performance. This model demonstrates the highest Area Under the Curve (AUC) in the ROC curve, indicating its strong classification capabilities and reliability in predicting churn.

NEXT STEPS

- **Variable Exploration:**
  - Analyze additional features such as internet bundles and other value-added services.
  - Assess the impact of these features on customer churn rates.
- **Understanding Churn Reasons:**
  - Conduct in-depth analysis to identify specific reasons for customer churn.
  - Utilize customer feedback and behavior data to pinpoint pain points and dissatisfaction causes.
- **Competitor Analysis and Strategy Development:**
  - Investigate competitors' offerings in the telecommunications industry.
  - Develop business strategies to enhance Syriatel's products and service offerings based on competitive insights.
- These steps will help Syriatel understand customer behavior better, reduce churn, and improve customer satisfaction by offering competitive and tailored services.

RECOMMENDATIONS

- Model Selection: Syriatel should adopt the Random Forest Classifier as the primary model for predicting customer churn. This model has shown excellent performance metrics, including high accuracy, F1-score, recall, and precision, making it highly effective for identifying customers likely to churn.
- Customer Retention Strategy: Syriatel should develop a customer retention strategy that focuses on key factors such as call minutes and charges. Personalized offers or discounts on day charges can be effective in retaining customers. By addressing these primary factors, Syriatel can enhance customer retention and reduce revenue loss.
- Customer Service Improvement: Since high volumes of customer service calls are closely linked to churn, Syriatel should implement strategies to minimize these calls. An advanced Interactive Voice Response (IVR) system could efficiently handle common issues, reducing the need for customer service calls and improving overall customer satisfaction.


THANK
YOU