Annolyne /
Phase3_Project

<> Code   Issues   Pull requests   Actions   Projects   Wiki   Security   Insights   Setting

☆ **0** stars   **0** forks   ⊙ **1** watching   **1** Branch   **0** Tags   Activity

🌐 Public repository

1 Branch   **0** Tags   🔍 Go to file   t   Go to file   +   Add file ▾   Code   ⋯

Annolyne fine tuned changes   5f1537c · 5 minutes ago   ⟲

Notebook.pdf	notebook.pdf	yesterday
Presentation.pdf	final commit	10 minutes ago
README.md	Update README.md	yesterday
bigml_59c28831336c6604c8...	First commit Data cleaning and EDA	4 days ago
data.txt	First commit Data cleaning and EDA	4 days ago
index.ipynb	Readme.md	yesterday

📖 **README**   ✏️ ☰

ML-Driven Churn Reduction: Syriatel's Customer Retention Strategy

Author: Annolyne Chemutai

School: Moringa school

Data Science PT

Phase 3 project

Project Overview

This project leverages the SyriaTel Telecoms dataset to construct a sophisticated classification model. The primary objective is to accurately forecast customer attrition, enabling SyriaTel to proactively identify and retain clients who may be considering termination of their services.

Business Understanding

In today's highly competitive telecommunications landscape, characterized by rapid technological advancements and diverse communication alternatives, it is crucial for Syriatel to leverage predictive analytics to effectively manage customer churn. Our objective is to build a robust predictive model that will empower Syriatel to implement targeted strategies to minimize churn, enhance customer retention, and expand its customer base, ultimately fostering sustained growth and profitability.

Key Benefits of the Predictive Model for Syriatel:

Churn Reduction: By accurately predicting customer churn, Syriatel can take proactive measures to retain customers, thereby increasing revenue and profitability.

Market Position Enhancement: Improved customer retention will strengthen Syriatel's market position, supporting growth and maintaining a competitive edge.

Enhanced Customer Experience: The model will facilitate more personalized and efficient customer service, leading to higher satisfaction and loyalty.

Increased Shareholder Value: A reduction in churn and growth in customer base will enhance revenue, driving up returns on investment (ROI) for shareholders.

Employee Growth and Compensation: As Syriatel's financial performance improves, employees will benefit from enhanced compensation, bonuses, and career development opportunities.

In summary, the adoption of this predictive model will not only optimize customer retention and satisfaction but also contribute to Syriatel's long-term success by boosting financial performance, market share, and employee rewards.

Dataset Overview

The SyriaTel Customer Churn dataset, sourced from Kaggle 'bigml_59c28831336c6604c800002a.csv' and located in the data directory, comprises 3,333 records with 21 features. It provides extensive details on telecom customers, including:

Customer Identification: State, account length, area code, and phone number.

Service Features: Availability of international and voicemail plans.

Usage Statistics:

Number of voicemails sent.

Total call metrics, including minutes and charges, segmented by day, evening, and night.

Metrics for international call usage.

Customer Service Interactions: Frequency of calls to customer service.

This dataset is instrumental for analyzing customer churn, offering insights into service usage patterns and customer behaviors critical for developing effective retention strategies.

Exploratory Data Analysis

Conducted a comprehensive exploration of the data through univariate, bivariate, and `multivariate analysis`.

This thorough data exploration aims to identify potential correlations among features and variable distributions, which will prove crucial for feature engineering and modeling.

Univariate Analysis

The dataset contains 3,333 customers, with 483 ending their contracts. This means 14.5% of customers were lost

Churn Distribution
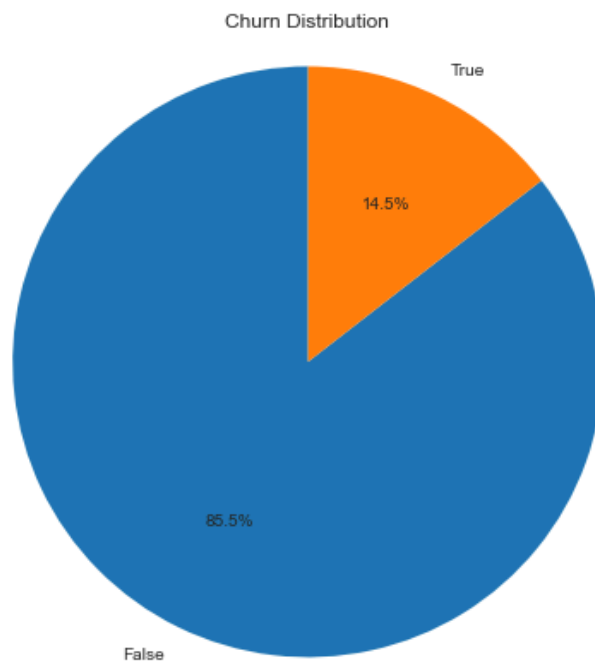
The uneven split between churned and non-churned customers creates a data imbalance. We need to fix this before modeling to avoid inaccurate predictions.

Bivariate Analysis

International plan subscribers had higher churn rates (42.4%) than non-subscribers (11.5%), suggesting potential issues with the plan. Conversely, voice mail plan subscribers showed lower churn rates (8.7%) compared to non-subscribers (16.7%), indicating it may reduce churn likelihood.
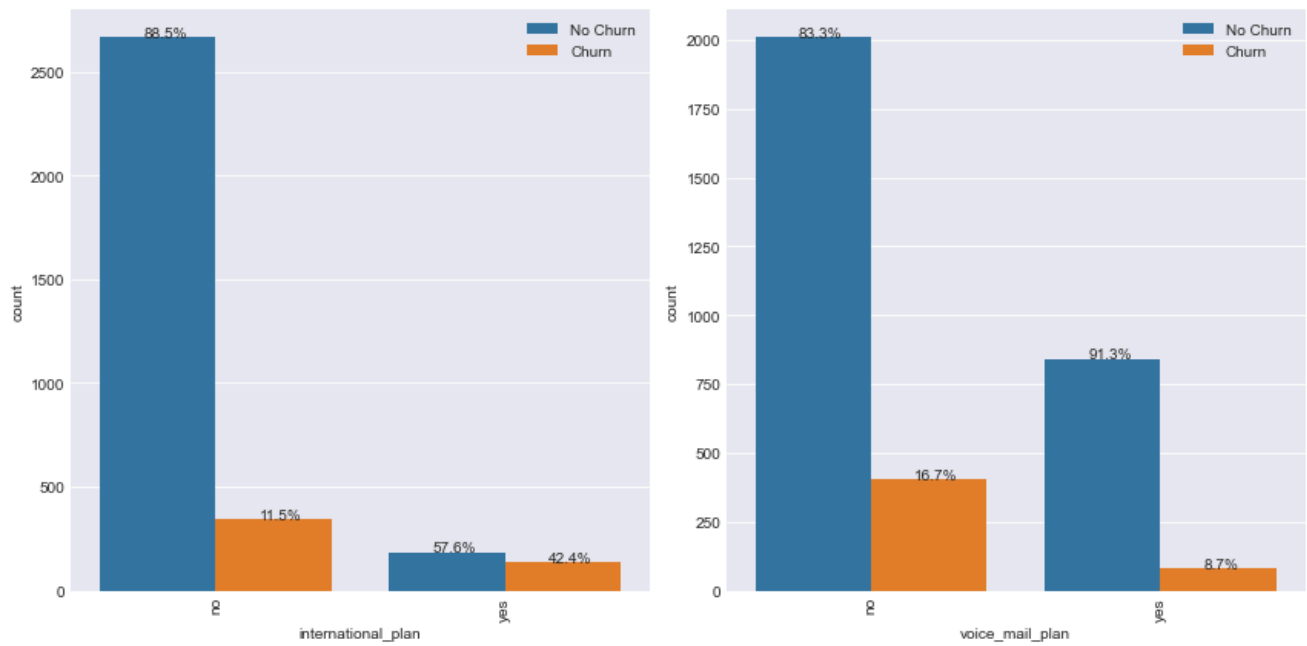
Feature correlations with customer churn using pairplots. This helps identify which factors may influence a customer's decision to leave.
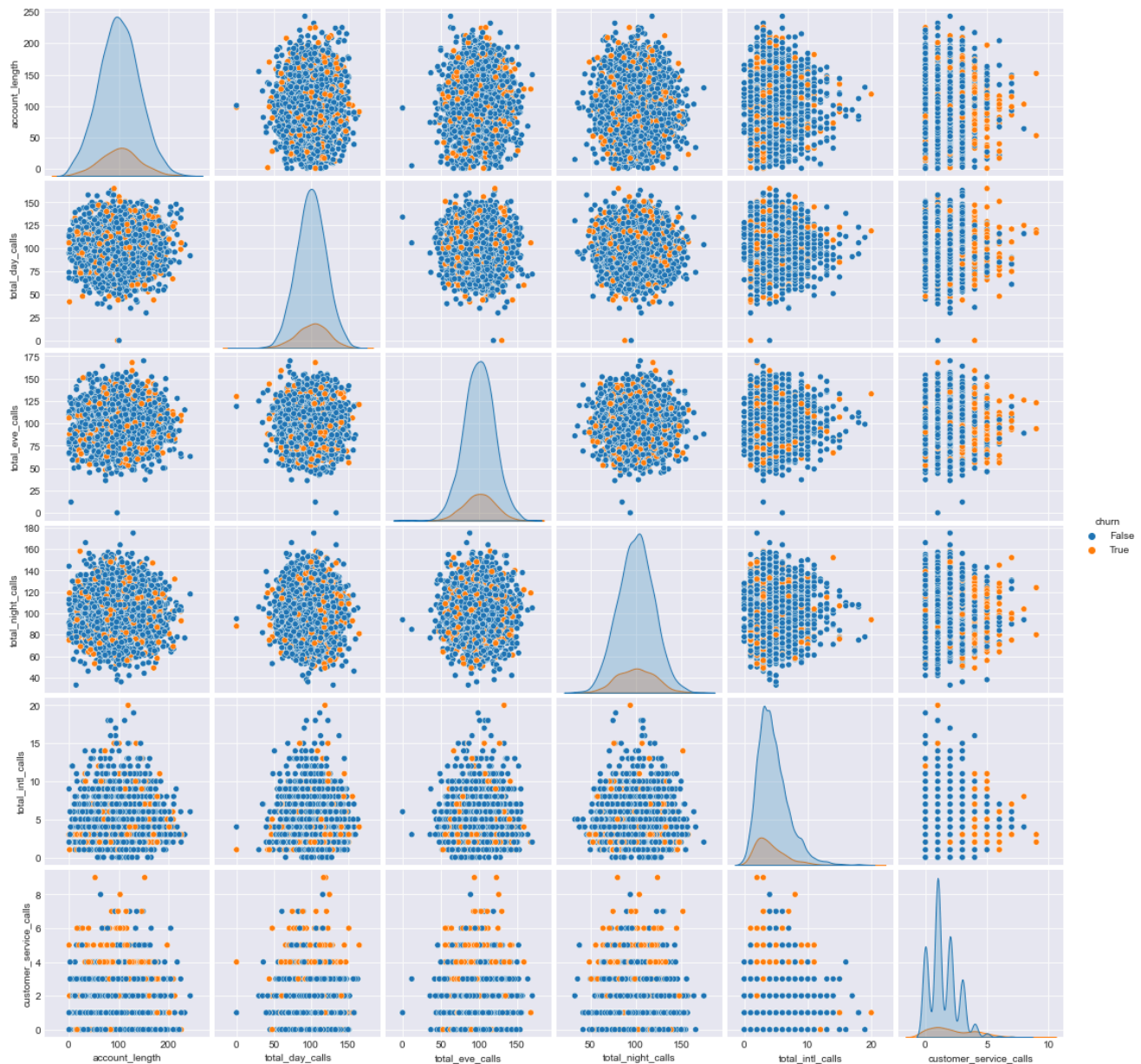
There is a significant correlation between the frequency of customer service calls and churn rates. The data indicates that customers who make more than four service calls are substantially more likely to terminate their service.

Furthermore, the high volume of customer service calls generally indicates dissatisfaction with the provided service. When customers need to make more than four calls, it suggests that their issues are not being resolved efficiently, which consequently increases the probability of service discontinuation..
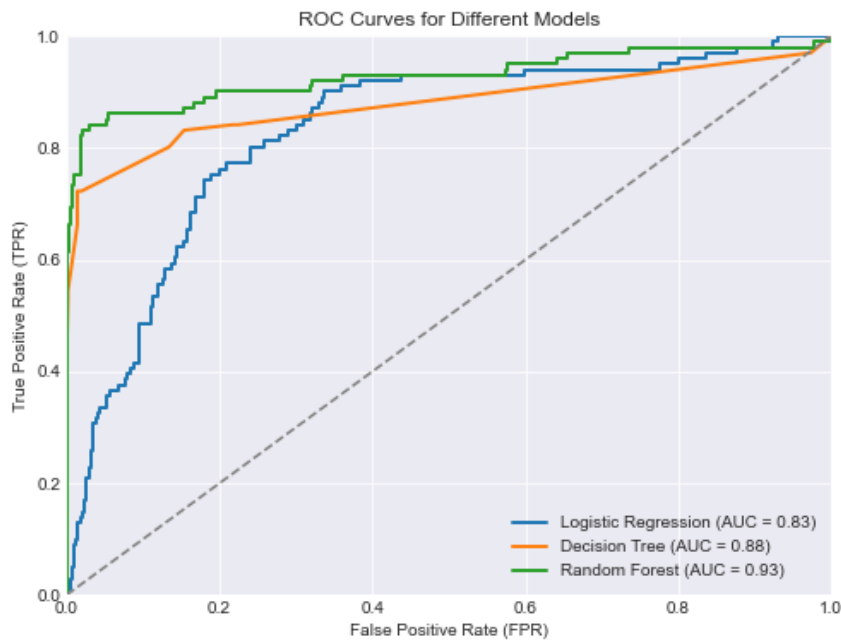
Modeling

Our analysis involved the development and evaluation of three distinct classification models to extract valuable insights and generate accurate predictions. The following provides a concise overview of each model employed:

Logistic Regression - AUC: 0.83

Decision Tree - AUC: 0.88

Random Forest - AUC: 0.93

Recommendation

Our predictive modeling and customer churn analysis have yielded valuable insights. Here are some actionable recommendations to boost customer retention at Syriatel Mobile Telecom:

1. Leverage the Best-Performing Model

Recommendation: Implement the **Random Forest model** for predicting customer churn. With the highest AUC (0.93), this model excels at distinguishing potential churners from loyal customers. Integrating it into your customer management system will enhance identification of high-risk clients.

Actions:

- **Develop a Churn Prediction System**: Deploy the Random Forest model in your CRM to score customers based on their churn likelihood.
- **Regular Monitoring**: Continuously assess and update the model to reflect evolving customer behavior and market trends.

Releases

No releases published
Create a new release

Packages

No packages published
Publish your first package

Languages

● **Jupyter Notebook** 100.0%