PREDICTIVE ANALYTICS FOR CUSTOMER CHURN MANAGEMENT

ENHANCING SYRIATEL'S TELECOMMUNICATIONS SERVICES

March, 2024

INTRODUCTION

Project Goals: Develop an effective predictive model for customer “churn” (discontinue services) in SyriaTel.

Overall Objective: Address revenue and reputation risks associated with churn.

Audience: Stakeholders in telecommunications industry.

BUSINESS UNDERSTANDING

Overview: SyriaTel faces customer churn challenges impacting revenue.

Specific Objectives: Identify churn patterns, predict customer behavior.

Stakeholders: Marketing, sales, customer service, management.

DATA UNDERSTANDING

DATASET: OBTAINED FROM KAGGLE, INCLUDES CUSTOMER DEMOGRAPHICS, USAGE, CHURN.

Data Preparation:
EDA, visualization,
correlation analysis.

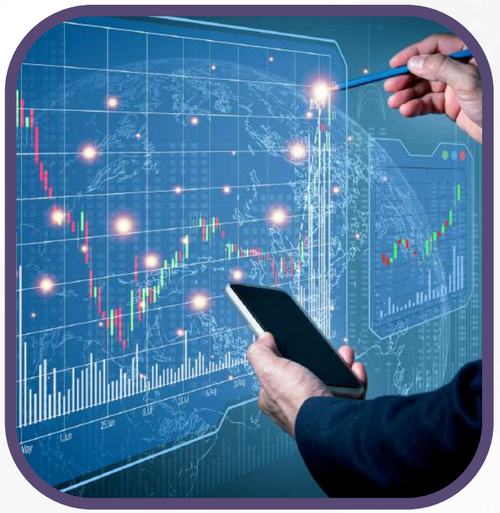
Attributes:
21 columns, 3333 rows,
target variable "churn".

DATA PREPARATION

EDA Techniques: Histograms, bar charts, correlation analysis.

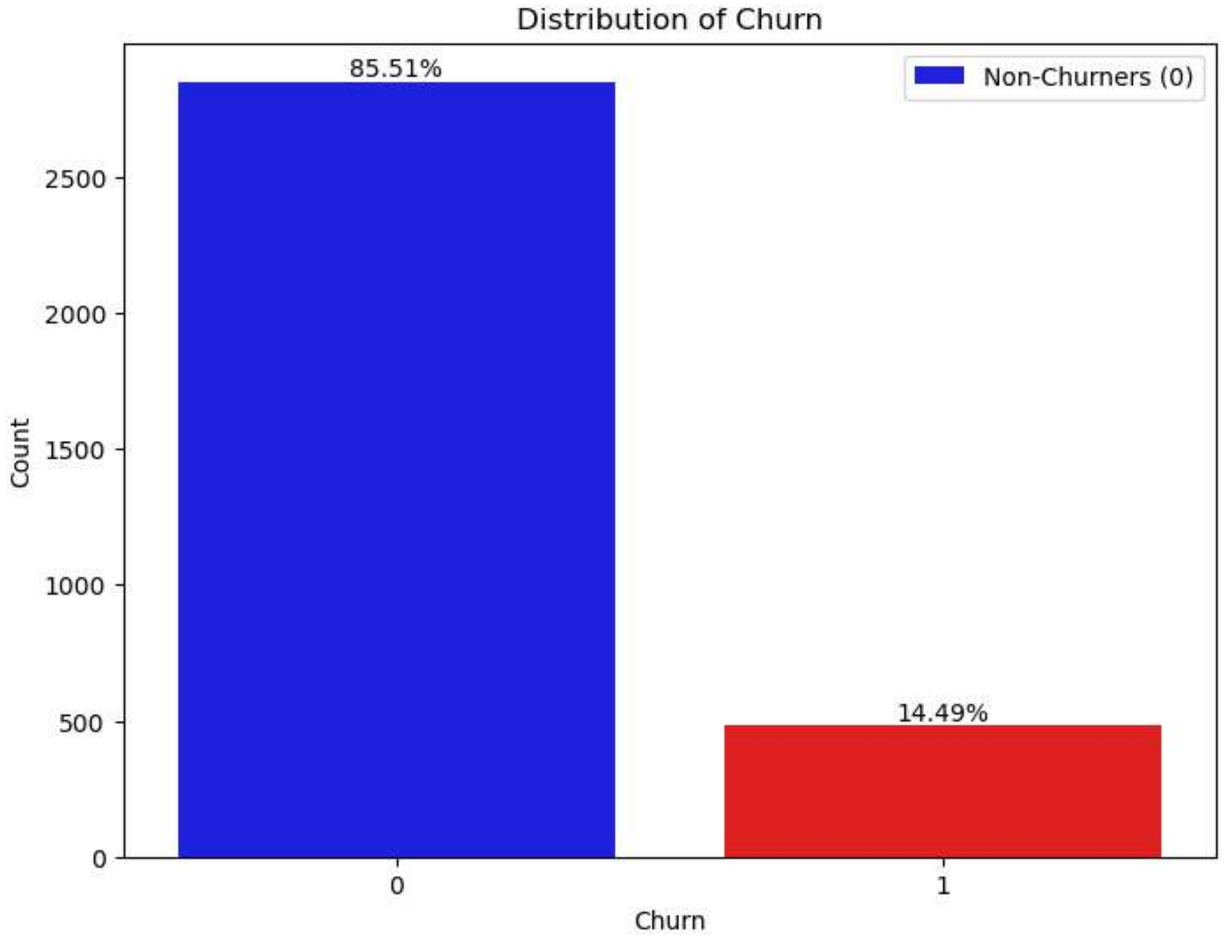
Data Visualization: Understanding variable distribution and relationships.

Feature Selection: Identifying influential predictors for churn.



EXPLORATORY DATA ANALYSIS (EDA)

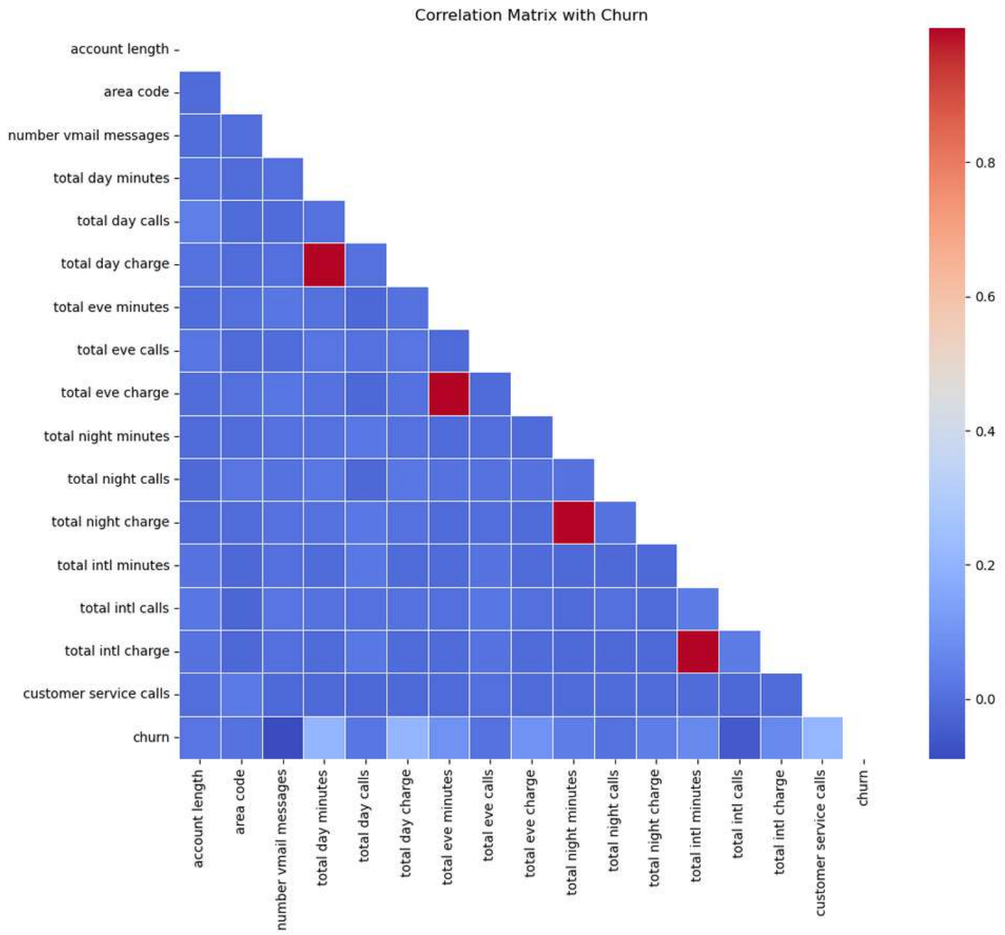
The graph represents the distribution of the target variable 'Churn'



Imbalanced Dataset: The dataset has a class imbalance, with churners comprising only 14% of the total records.

EXPLORATORY DATA ANALYSIS (EDA)

The graph represents the Correlation Matrix of the numerical variables and the target variable 'Churn'



Based on the correlation analysis,
the most influential features in predicting churn:

* Customer service calls.
* Total day minutes.
* Total day charge.
* Total eve minutes.

MODELING

MODEL DEVELOPMENT: 4 MODELS WERE DEVELOPED

01

Baseline Logistic regression Model.

02

Random Forest Model

03

Tuned Random Forest Model.

04

XGBoost Model.

MODELING

BASELINE LOGISTIC REGRESSION MODEL.

- Accuracy: The accuracy of the model was approximately 82.91%.

  A measure of the overall effectiveness of the churn prediction model in correctly classifying customers as churners or non-churners.

- Precision: The precision of the model was approximately 46.45%.

  A measure of the accuracy of the model in identifying customers who are likely to churn “Discontinue services” with SyriaTel.

MODELING

MORE-COMPLEX MODEL - RANDOM FOREST MODEL

The performance metrics for the Random Forest Model were much better than those of the Baseline Model

- Accuracy: The accuracy of the model was 96.10%.
- Precision: The precision of the model was 93.10% .

MODELING

TUNED RANDOM FOREST MODEL

The Tuned Random Forest Model did not perform any better than the original Random Forest Model, but it was still better than the Baseline Logistic Regression model.

- Accuracy: The accuracy of the model was 95.20%.
- Precision: The precision of the model was 86.30%.

.

MODELING

THE XGBOOST MODEL

The XGBoost model performed the best out of all the models in terms of performance metrics such as accuracy and precision.

- Accuracy: The accuracy of the model was 96.40%.
- Precision: The precision of the model was 95.30%.

.

MODEL EVALUATION

EVALUATING THE MODELS BASED ON THE PERFORMANCE METRICS

1. General performance metrics of accuracy, precision, recall, and F1 score.

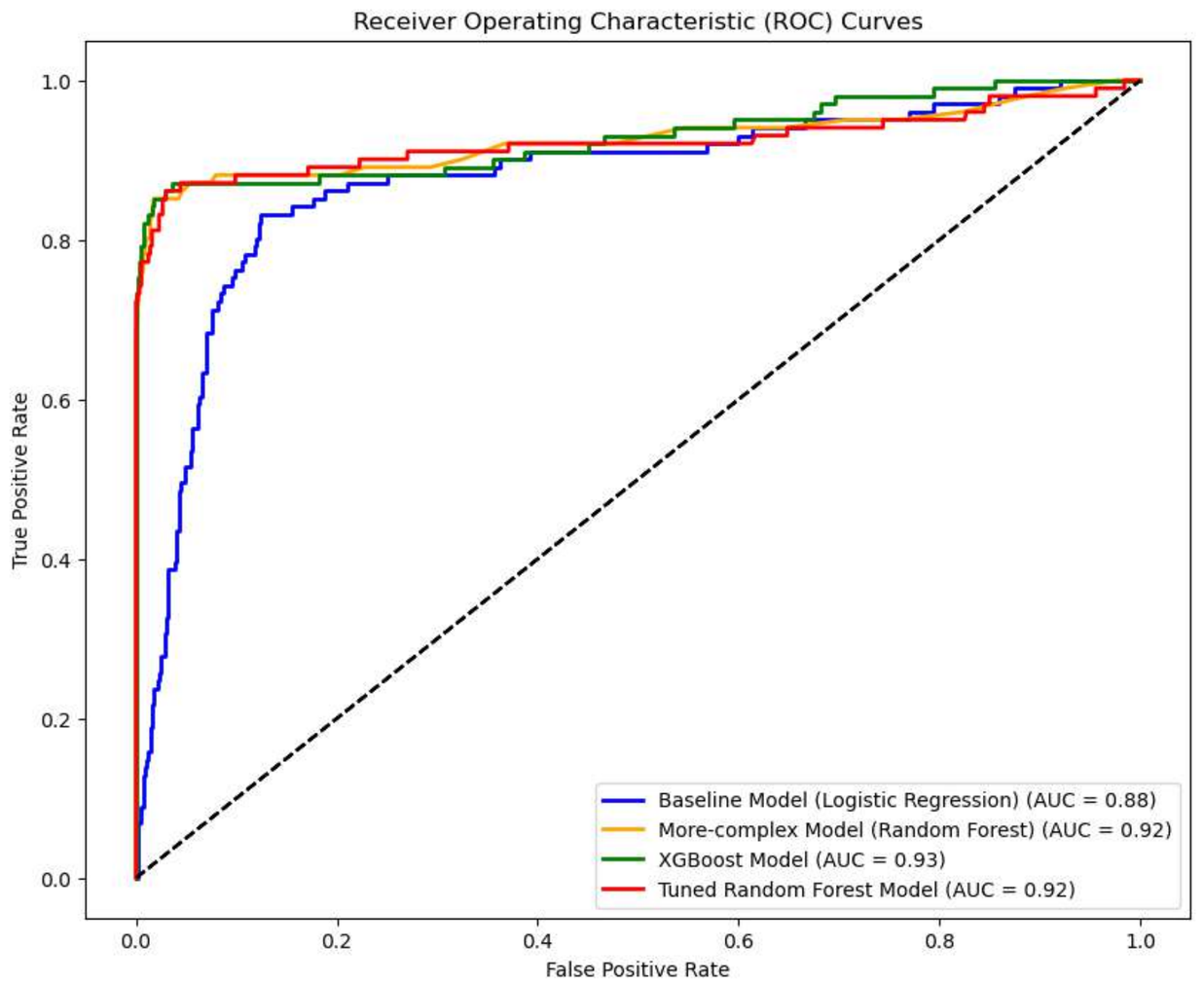
- The XGBoost Model performed the best among all the models in this measure.

2. The ROC AUC (Receiver Operating Characteristic - Area Under the Curve) metric

- A measure of how well a model distinguishes between churned and non-churned customers.

MODEL EVALUATION

THE ROC AUC (RECEIVER OPERATING CHARACTERISTIC - AREA UNDER THE CURVE) METRIC



Performance Rank:

1. The XGBoost Model: ROC AUC Score of 0.93
2. The Random Forest Models: 0.924
3. The Tuned Random Forest: 0.922
4. The Baseline Model: 0.88.

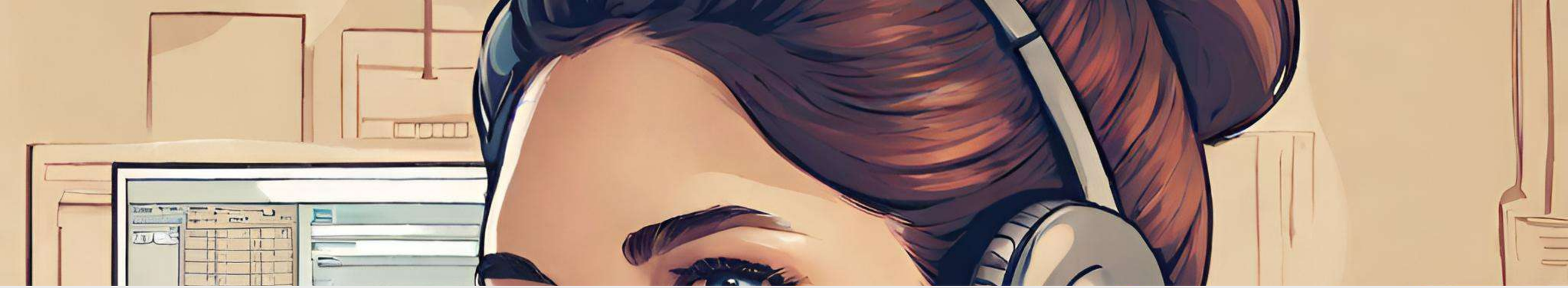
- The XGBoost Model performed the best among all the models in the ROC AUC measure.

RECOMMENDATIONS

- TARGETED RETENTION STRATEGIES: PERSONALIZED OFFERS, PROACTIVE INTERVENTIONS.
- CUSTOMER EXPERIENCE ENHANCEMENT: SERVICE QUALITY IMPROVEMENTS, ADDRESSING PAIN POINTS.
- CONTINUOUS MONITORING: MODEL PERFORMANCE ASSESSMENT, ADAPTATION.

NEXT STEPS

- DEEPER ANALYSIS: EXPLORE ADDITIONAL DATA SOURCES, DEMOGRAPHIC FACTORS.
- CONTINUOUS IMPROVEMENT: REFINE MODELS, ADAPT STRATEGIES FOR EVOLVING TRENDS.



THANK YOU