PREDICTIVE ANALYTICS FOR CUSTOMER CHURN MANAGEMENT

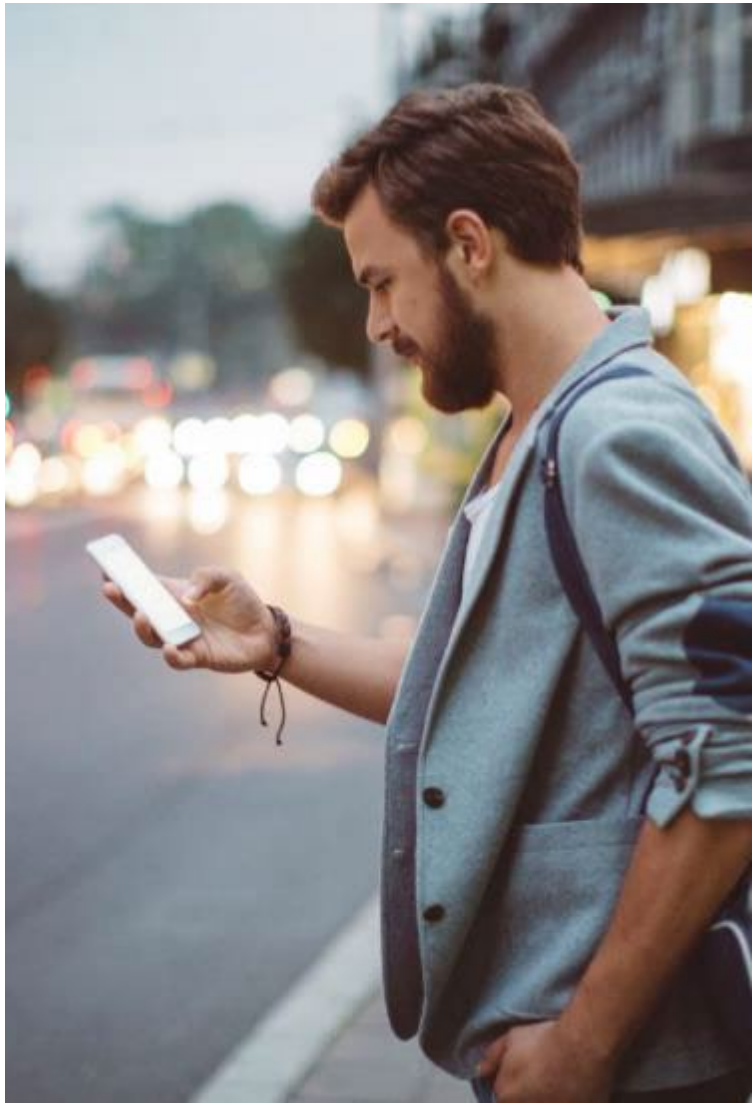
ENHANCING SYRIATEL'S TELECOMMUNICATIONS SERVICES

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

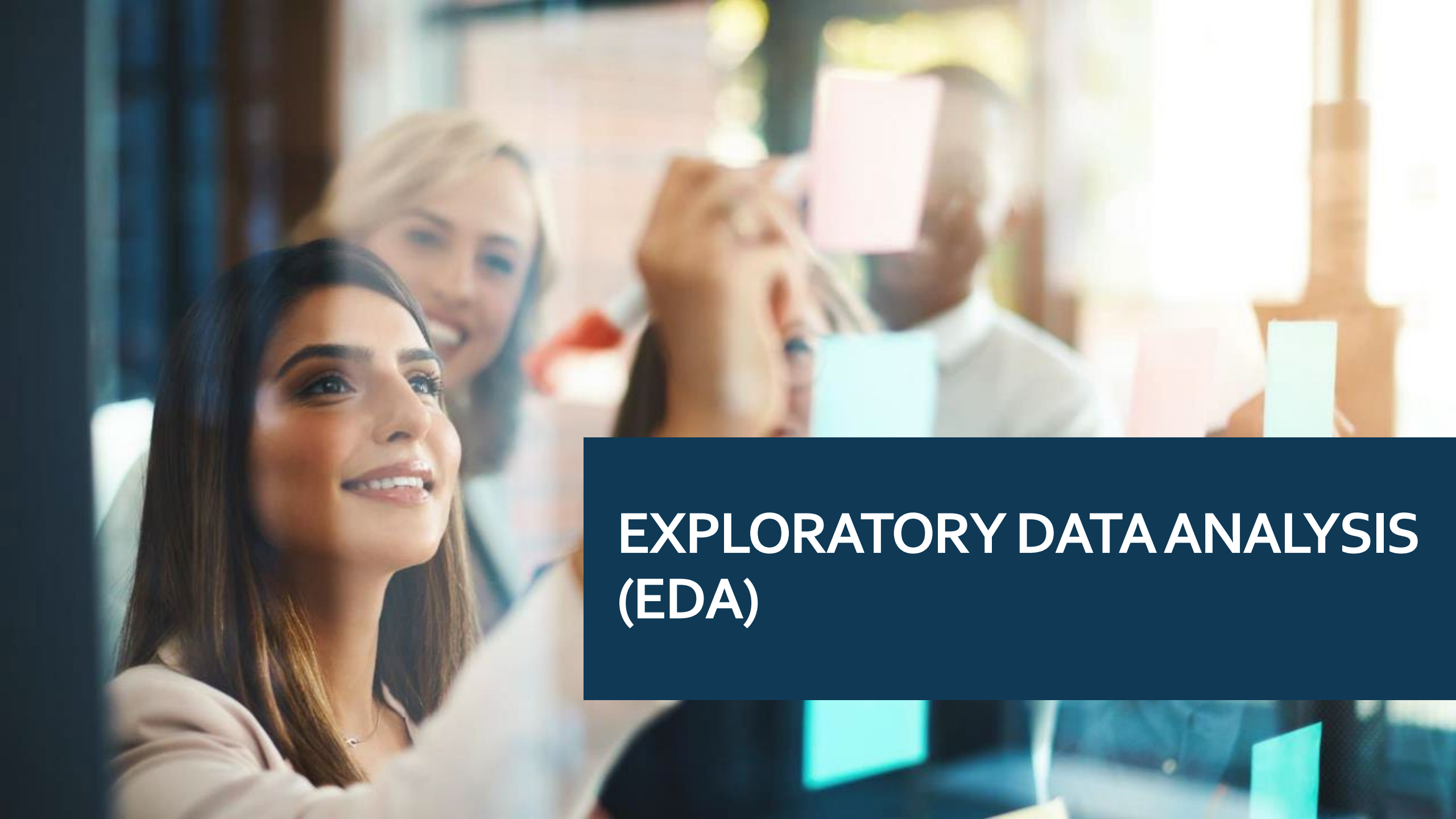
DATASET: OBTAINED FROM KAGGLE, INCLUDES CUSTOMER DEMOGRAPHICS, USAGE, CHURN.

Data Preparation: EDA, visualization, correlation analysis.

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

DATA PREPARATION

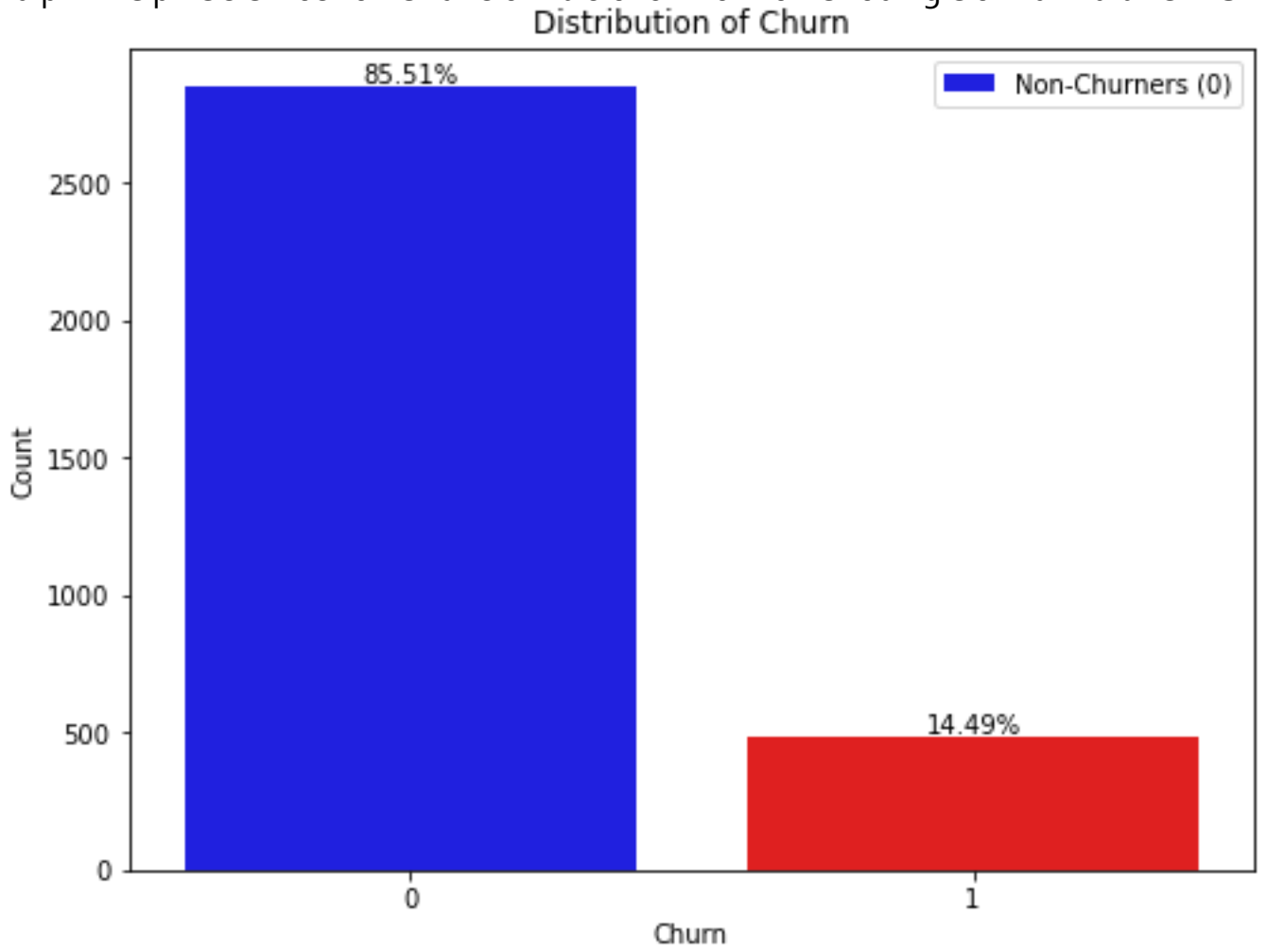
- **EDA Techniques**: Histograms, bar charts, correlation analysis.
- **Data Visualization**: Understanding variable distribution and relationships.
- **Feature Selection**: Identifying influential predictors for churn.



EXPLORATORY DATA ANALYSIS (EDA)

EXPLORATORY DATA ANALYSIS (EDA)

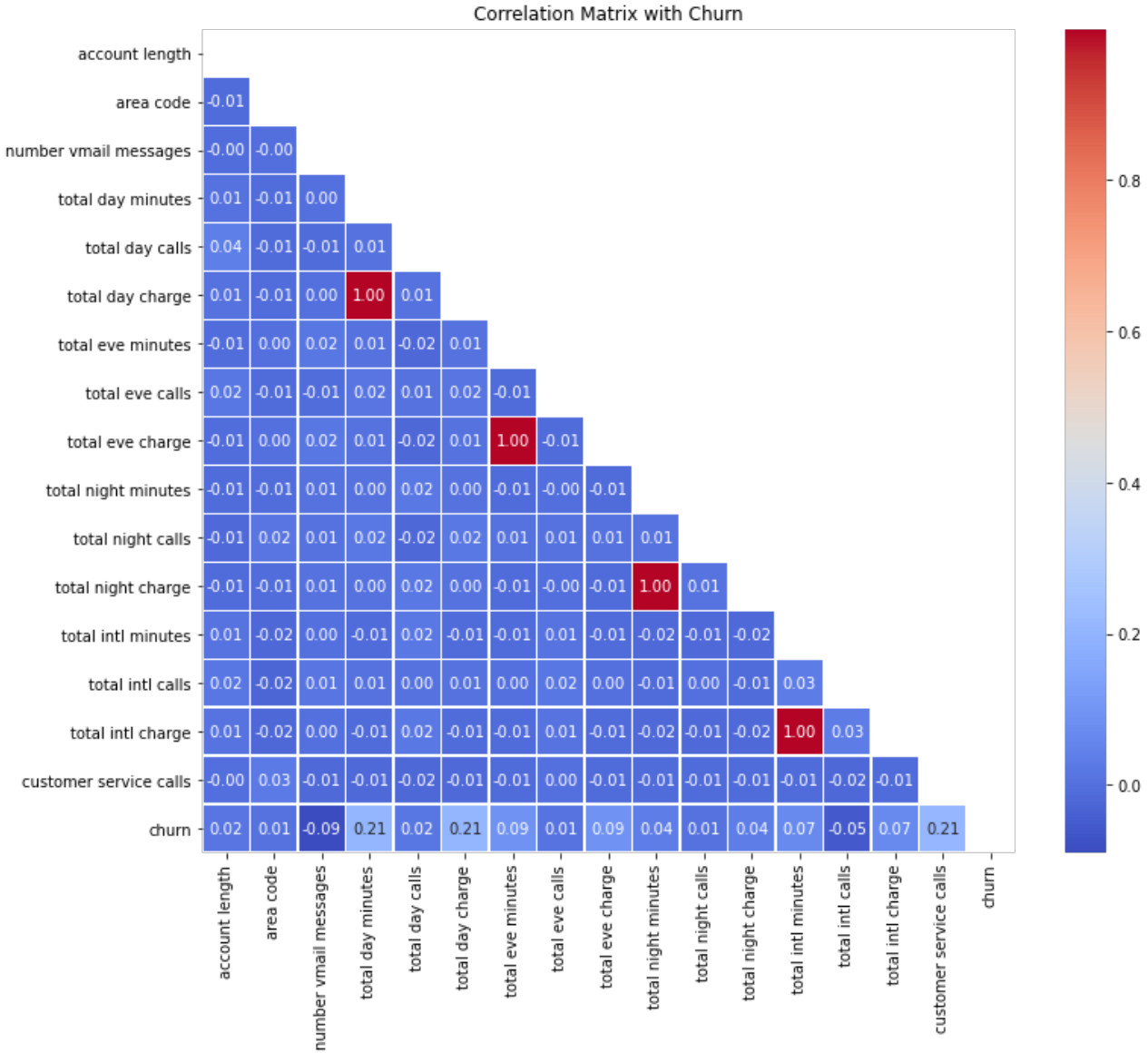
The graph represents the distribution of the target variable 'Churn'



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

MODELLING

Model Development.

3 models were developed.

1. Baseline Logistic regression Model
2. Decision Tree Model
3. Random Forest Model

MODELING

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.

BASELINE LOGISTIC REGRESSION MODEL

MODELING

DECISION TREE MODEL

The performance metrics for the decision tree model were better to those of the baseline model.

Accuracy: The accuracy of the model was approximately 91.45%.

Precision: The precision of the model was approximately 69.64%.

MODELING

RANDOM FOREST MODEL

The performance metrics for the random forest model were better to those of the decision tree model.

Accuracy: The accuracy of the model was approximately 96.1%.

Precision: The precision of the model was approximately 93.1%.

MODELLING

Model Evaluation.

EVALUATING THE MODELS BASED ON THE PERFORMANCE METRICS

MODEL EVALUATION

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

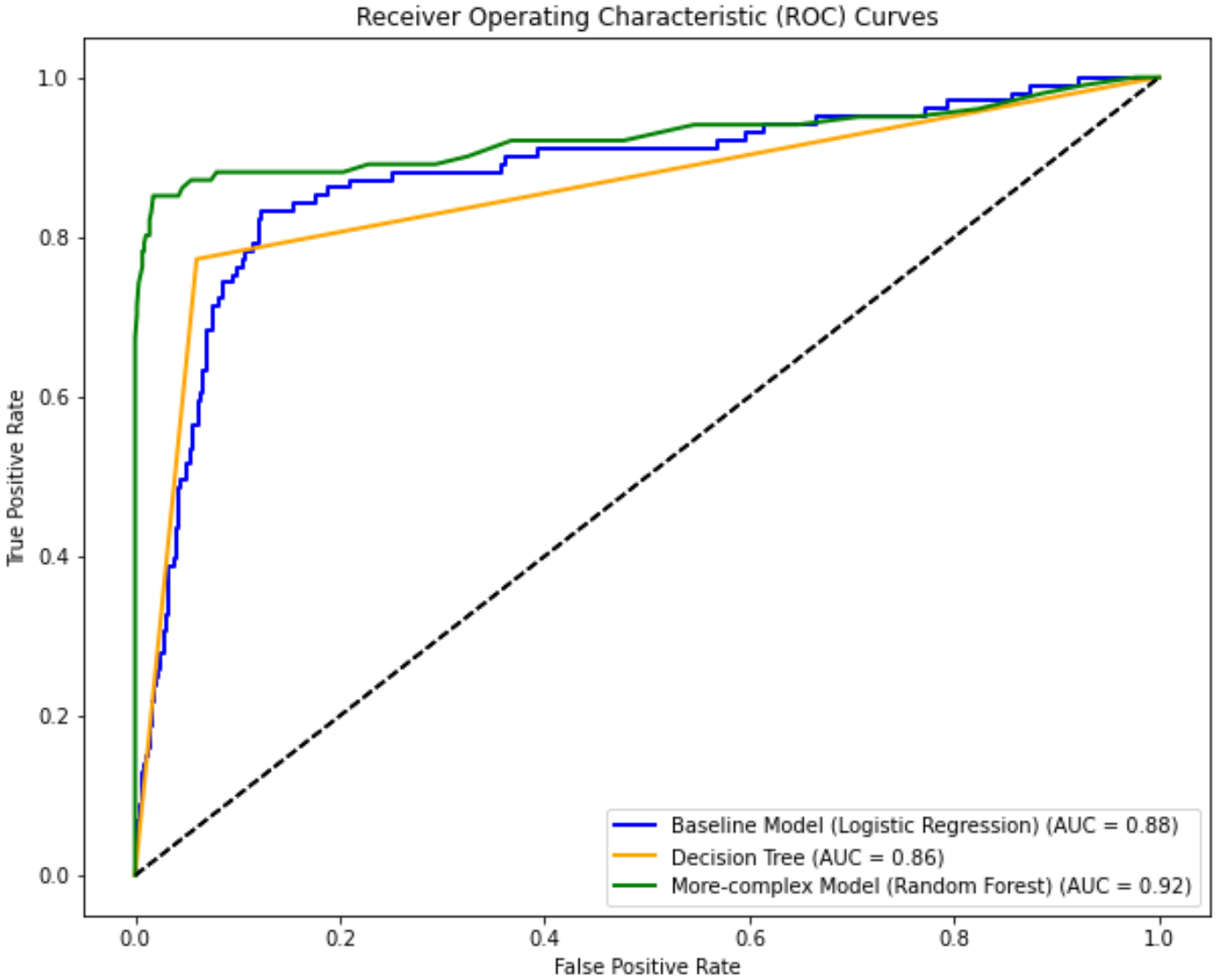
1. The Random Forest Model performed the best among all the models in this measure.

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 Random Forest Model: ROC AUC Score of 0.92
2. The Decision Tree Model: 0.86
3. The Baseline Model: 0.88.

The Random Forest 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.