Churn Prediction Modeling:

# Predicting customer churn in SyriaTel

# Introduction

**Overview**

- Developing models to predict customer churn for SyriaTel, a telecommunications company. The goal is to assist the telecom industry in minimizing revenue loss by identifying patterns indicative of customers likely to leave soon.

**Business problem**

- Targeting SyriaTel's business needs, this project aims to optimize customer retention strategies by analyzing predictable patterns, benefiting the company's financial sustainability.

# Dataset Overview

- The data used in this project is from SyriaTel telecommunications company: https://www.kaggle.com/datasets/becksddf/churn-in-telecoms-dataset.
- Variables: some columns in the data are, Customer service calls, total day calls, total eve calls, total intl calls, total night calls, total day minutes etc.
- Target variable: churn
- We will use this information to predict the future churn and give insights into what can be done to reduce the churn rate by determining what contributes to churn.

# Exploratory Data Analysis

- Data is clean.
- Visualizing Data Distributions to see correlations of various variables.
- Identifying Correlations of all variables to churn.
- Encoding categorical variables.

# Feature Engineering

- Created new features Total Call Minutes, Total Call Charges, Average Call Duration, usage patterns, time-of-day patterns, combined features and created binary features.
- Significance of Engineered Features: to improve model predictiveness.

# Model Selection

Applied the below models:

- Baseline Model: Logistic Regression
- Complex Model: Random Forest
- Hyperparameter Tuning and Optimization

# Model Evaluation

Model Performance

- Metrics: Accuracy, Precision, Recall, F1 Score, ROC AUC

Baseline Model Metrics: Accuracy: 0.85 Precision: 0.67 Recall: 0.06 F1 Score: 0.11 ROC AUC: 0.53

Complex Model Metrics: Accuracy: 0.97 Precision: 0.97 Recall: 0.83 F1 Score: 0.89 ROC AUC: 0.91

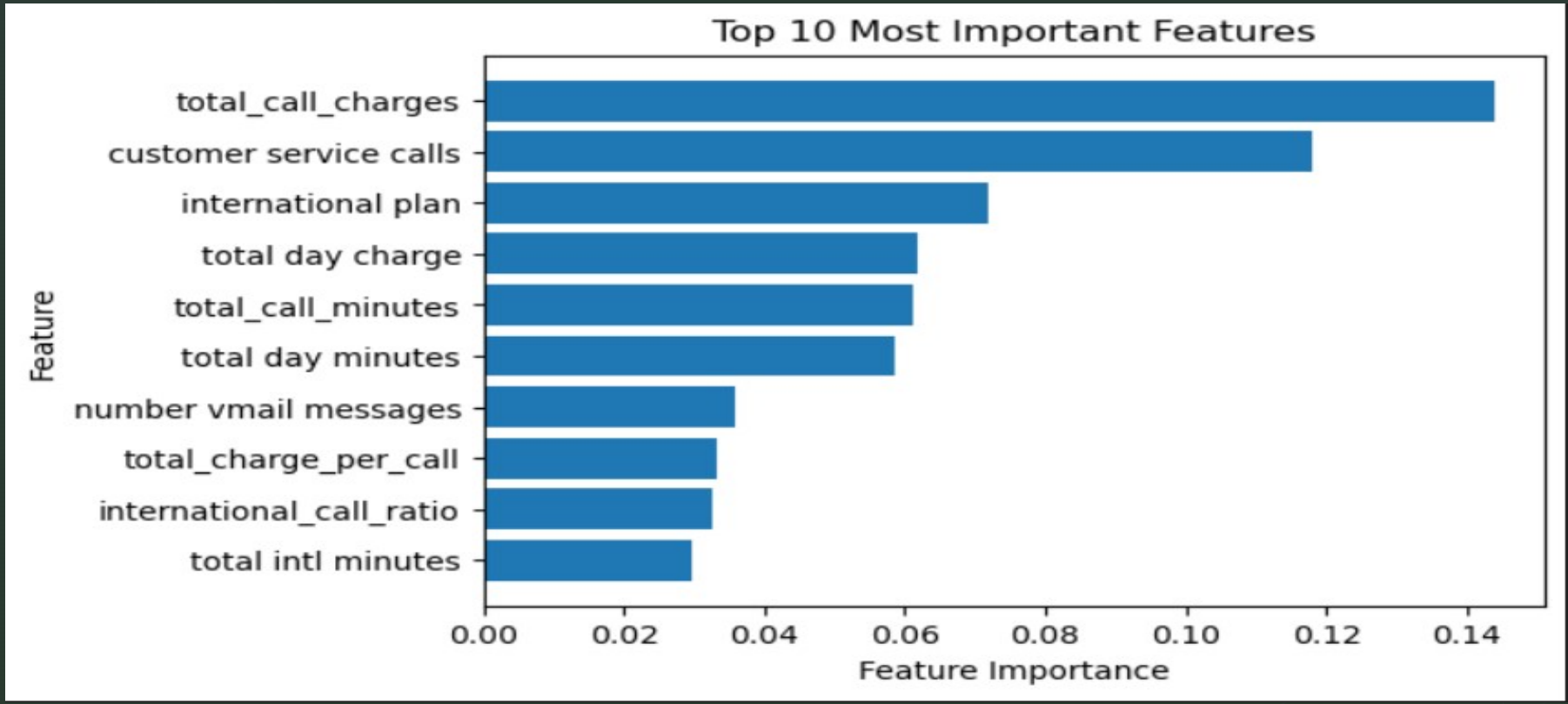
Tuned Model Metrics: Accuracy: 0.97 Precision: 0.97 Recall: 0.83 F1 Score: 0.89 ROC AUC: 0.91

- Comparison of Baseline, Complex, and Tuned Models

Baseline model performance is not better than random chance while complex model and tuned model show a strong ability to distinguish variables that affect churn

# Feature importance

- Features most important in predicting churn.



# Conclusion

The model was useful in below scenarios:

- Customers with high usage patterns give a more accurate prediction for customers who are unlikely to churn.
- Customers with shorter contracts are likely to predict churn rate more accurately.
- Customers with customer service interactions indicate customer dissatisfaction hence may predict likelihood of churning.
- Customer segmentation: factors like age, location influence churn patterns, if specific segments have distinct churn behaviour ,the model can be useful in predicting churn within these segments.

# Business recommendations

- Focus marketing efforts on customers identified as high risk for churn.
- Enhancing customer support to reduce customer service calls.
- Product/service enhancements that is the ones that are most popular among long-term customers.
- Establish customer feedback channels to understand reasons behind customer churn.

# Limitations

- Constant change of customer behaviour would not reflect future behaviour clearly.
- Unseen patterns not reflected in the data may not present full picture of customer churn reasons.
- Incomplete/Limited features or inaccurate data may give wrong or insufficient insights.

# THANK YOU!