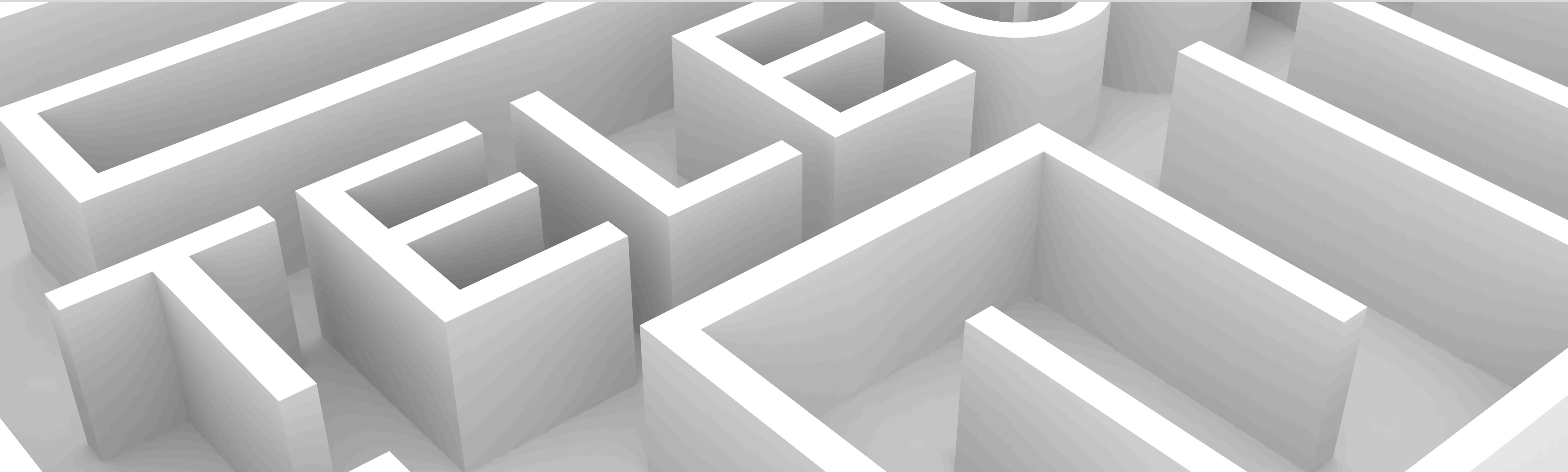
# SyriaTel: Reducing Customer Churn Through Predictive Analysis

A Data-Driven Future

# Project overview

Using historical data and machine learning to predict and mitigate customer churn



# Business context and Understanding

**Churn** : refers to customer attrition where individuals stop using a service or product

**Importance**

- Causes loss of revenue, market share, and brand reputation

**Strategic goal**

- Proactively identify high-risk customers and mitigate attrition

# Project Objectives

**Main Objective :**

- Predict which customers are likely to churn.

**Specific objectives:**

- Understand key drivers behind customer churn.
- Inform targeted retention strategies

# Data understanding

**Data description**

- Scope: Customer usage patterns, subscriptions, and customer service interaction
- Account-Level: Account length.
- Subscription Plans: International plan and voice mail plan.
- Day usage: total day minutes, day calls, total day charge
- Evening usage: total evening minutes, evening calls, evening charge
- Night usage: total night minutes, night calls, night charge
- International usage: total international minutes, calls, charge
- Support Interactions: Number of customer service calls.
- Target Variable: Churn (1 = yes, 0 = no)

# Data pre-processing

- Dropped irrelevant column (phone number)
- one hot encoded categorical features and label encoded the target
- Standardized the numerical features
- Split the data into training (80%) and testing (20%)
- Addressed class imbalance by using SMOTE( Synthetic Minority oversampling Technique)

# Modelling approach

Evaluated 4 models

1. Logistic regression
2. Decision tree
3. LGBoost
4. Random Forest

# Model Performance evaluation

Evaluation Metrics used

- Accuracy for overall correctness of prediction
- Confusion metric for a detailed breakdown of predictions
- Precision, Recall, and F1 score on churn detection
- ROC-AUC for model discrimination ability
- Feature importance

# Results and insights

**Logistic Regression Model Evaluation**

---

Accuracy: 0.7886056971514243

---

Confusion Matrix:

```
[[450 116]
 [ 25  76]]
```

---

Classification Report:

| | precision | recall | f1-score | support |
|---|---|---|---|---|
| 0 | 0.95 | 0.80 | 0.86 | 566 |
| 1 | 0.40 | 0.75 | 0.52 | 101 |
| accuracy | | | 0.79 | 667 |
| macro avg | 0.67 | 0.77 | 0.69 | 667 |
| weighted avg | 0.86 | 0.79 | 0.81 | 667 |

---

ROC AUC Score: 0.8307560438022599

---

**Decision Tree Results**

---

Accuracy: 0.9385307346326837

---

Confusion Matrix:

```
[[558   8]
 [ 33  68]]
```

---

Classification Report:

| | precision | recall | f1-score | support |
|---|---|---|---|---|
| 0 | 0.94 | 0.99 | 0.96 | 566 |
| 1 | 0.89 | 0.67 | 0.77 | 101 |
| accuracy | | | 0.94 | 667 |
| macro avg | 0.92 | 0.83 | 0.87 | 667 |
| weighted avg | 0.94 | 0.94 | 0.93 | 667 |

---

ROC AUC Score: 0.8295665255571495

---

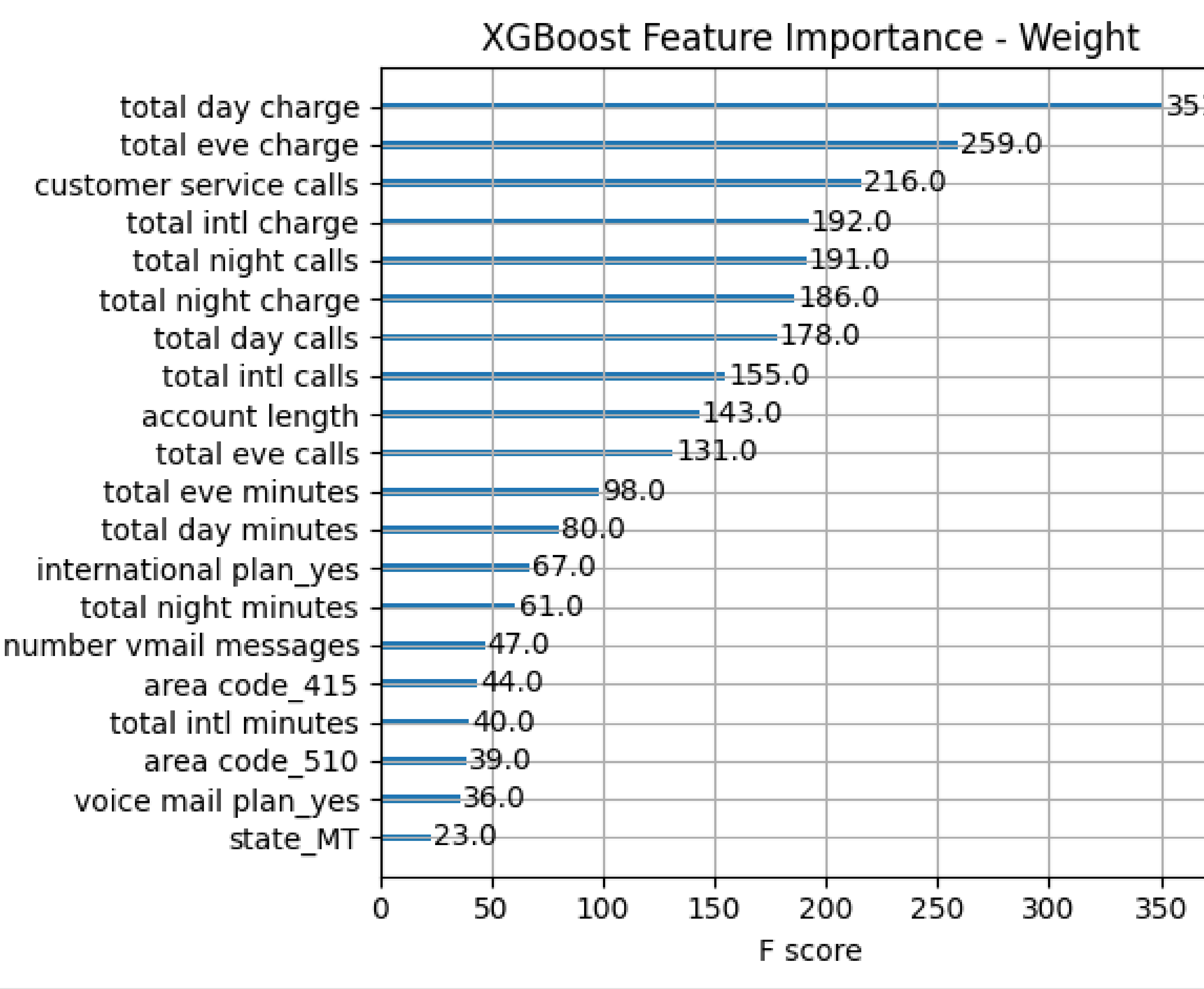
# Results and insights

```
                XGBoost Results
-----------------------------------------------------
       Accuracy: 0.9580209895052474
-----------------------------------------------------
             Confusion Matrix:
                 [[558   8]
                  [ 20  81]]
-----------------------------------------------------
            Classification Report:
              precision    recall  f1-score   support

           0       0.97      0.99      0.98       566
           1       0.91      0.80      0.85       101

    accuracy                           0.96       667
   macro avg       0.94      0.89      0.91       667
weighted avg       0.96      0.96      0.96       667

-----------------------------------------------------
        ROC AUC Score: 0.8939229612007138
-----------------------------------------------------
```



# Recommendations

- The XGBoost model had the overall best performance
- High feature importance in

1. Customer service calls: Need to focus on the quality of customer service
2. Usage metrics evening, day charges: need to consider incentives

# Next steps

- Hyperparameter tuning for the Decision Tree, Random Forest, and XGBoost

- Further analysis of customer feedback to assist in developing targeted churn mitigation interventions

# Thank you!