



# **Customer Churn Prediction Model – SyriaTel**

# Business Overview

## **Objective:**

To build a model that predicts whether a customer is likely to churn (stop using SyriaTel services).

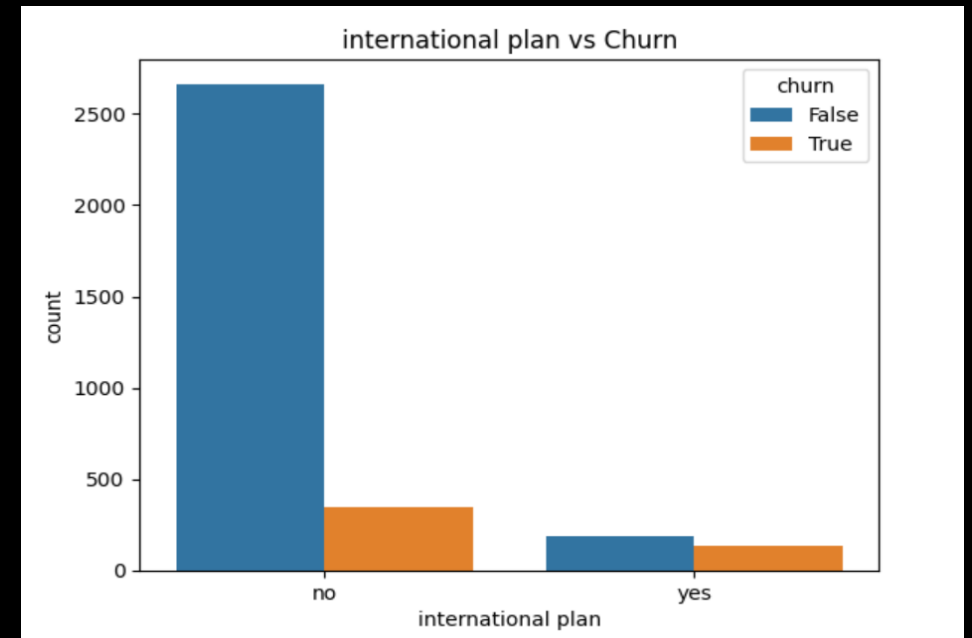
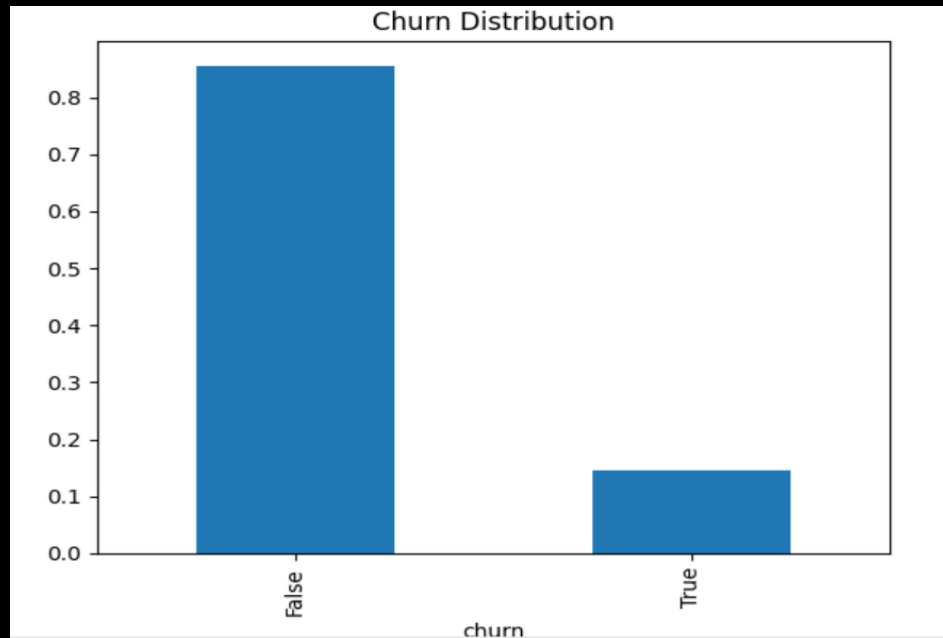
## **Why it matters:**

- Acquiring new customers is 5x more expensive than retaining existing ones.
- Early detection of churn risk helps increase customer loyalty and revenue.
- Aligns with SyriaTel's goal of enhancing customer satisfaction and retention.

# Exploratory Data Analysis (EDA)

## What I Explored:

- **Demographics:** Customer location and phone usage
- **Behavioral data:** Call duration, number of calls, service plans
- **Churn rate:** ~10% of customers churned



## Key Observations:

- High churn linked to multiple customer service calls
- Customers without voice mail or international plans are more likely to churn
- Usage patterns (day vs night) provide predictive signals

# Data Processing and Modeling

## How I Built the Model:

### 1.Data Cleaning

- Removed irrelevant fields (e.g., phone number)
- Converted text to numbers (e.g., Yes/No to 1/0)

### 2.Data Balancing

Only 10% churned → I used the SMOTE technique to balance data. This is an important step because if not undertaken, classification models tend to:

- Favor the majority class (non-churners).
- Show high accuracy, but miss the minority class (actual churners).
- Make poor predictions for the very thing you're trying to detect

### 1.Model Training

I Tried various algorithms; then chose one with highest accuracy and reliability Random Forest Classifier was likely chosen due to High ROC-AUC score (~0.917).

The Gradient Boosting Machine had a slightly lower accuracy of ROC-AUC of 0.912 compared to this. After the tuning of the model, there was just a slight improvement in terms of change.

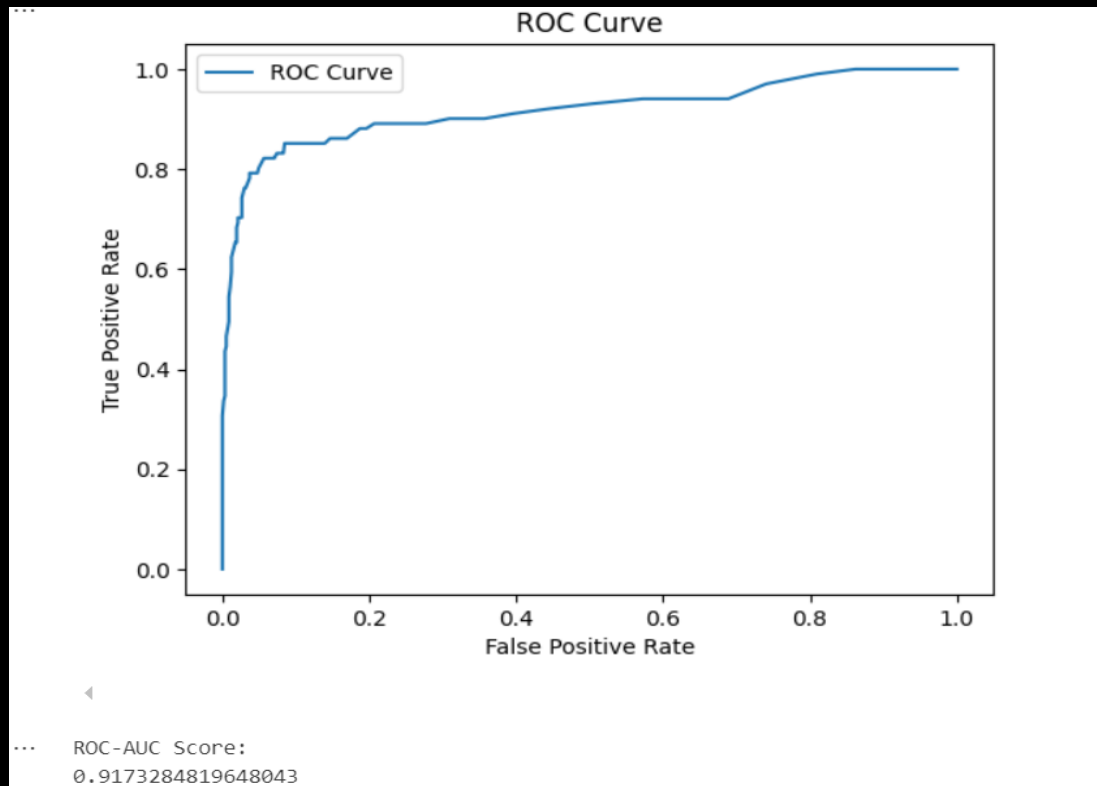
### Outcome:

Built a model that learns from customer behavior to predict churn before it happens.

# Model Performance & Results

**Metric Used:**

**ROC-AUC Score = 0.917** (Excellent Performance)



Classification Report:

	precision	recall	f1-score	support
False	0.96	0.95	0.96	566
True	0.74	0.80	0.77	101
accuracy			0.93	667
macro avg	0.85	0.88	0.86	667
weighted avg	0.93	0.93	0.93	667

Confusion Matrix:

```
[[538 28]
 [ 20 81]]
```

Accuracy Score:

0.9280359820089955

**Interpretation:**

- **91.7% accuracy** in distinguishing between churners and non-churners.
- The model is reliable and can be used for real-time customer scoring.

# Key Insights

## High Churn Risk Detected

The model shows high accuracy ( $AUC = 0.917$ ) in the churn prediction.

This means that certain customer behaviors strongly correlate with churn, such as:

- (a) High number of customer service calls.
- (b) Not subscribing to voice mail or international plans.
- (c) High charges during daytime usage.

# Recommendations for the Business

## **1.Proactive Retention:**

1. Call or offer discounts to predicted high-risk customers

## **2.Enhance Customer Support:**

1. Improve issue resolution speed and quality

## **3.Bundle Plans & Offers:**

1. Encourage uptake of voicemail and international packages

## **4.Data-Driven Strategy:**

1. Use the model monthly to flag customers needing attention

# Conclusion & Next Steps

## Conclusion:

The model provides early warning signals to prevent churn and save costs.

## Next Steps:

- Deploy the model into SyriaTel's CRM system.
- Continuously monitor and update with fresh data.
- Expand this approach to other areas (upselling, segmentation).





**Presented by:**

Kelvin Shilisia

Date: 22<sup>nd</sup> July 2025

LinkedIn: [www.linkedin.com/in/kelvin-shilisia-2b289b108](https://www.linkedin.com/in/kelvin-shilisia-2b289b108)

Email: kelvinshilisia@gmail.com