

Brucethegreat4 /  
phase3-syriaTel\_customer\_churn

&lt;&gt; Code

Issues

Pull requests

Actions

Projects

Wiki

Security

In

phase3-syriaTel\_customer\_churn

/ README.md in main

Cancel changes

Commit changes...

Edit

Preview

☐ Show Diff

# phase3-syriaTel\_customer\_churn

## Project Overview

This project focuses on predicting customer churn for SyriaTel, a telecommunications company, using machine learning. By identifying at-risk customers, the company can implement effective retention strategies, improve customer satisfaction and reduce revenue loss. The project involves exploratory data analysis (EDA), data preprocessing, model development and actionable recommendations to address churn.

## Business Understanding

Customer churn is a critical issue for SyriaTel, as it directly impacts profitability and growth. Retaining existing customers is more cost-effective than acquiring new ones, making churn prediction a high-priority business problem. By use of machine learning, SyriaTel can proactively identify customers likely to churn and take measures to retain them.

### Problem Statement:

- SyriaTel is experiencing customer churn, leading to revenue loss and increased operational costs. The company lacks a systematic approach to predict and address churn. The goal is to build a robust predictive model that identifies customers at risk of churning, enabling SyriaTel to implement targeted retention strategies and reduce churn rates.

### Objectives:

1. Develop a classification model to predict customer churn.
2. Identify the key drivers influencing churn.
3. Provide actionable insights for customer retention and business strategy.

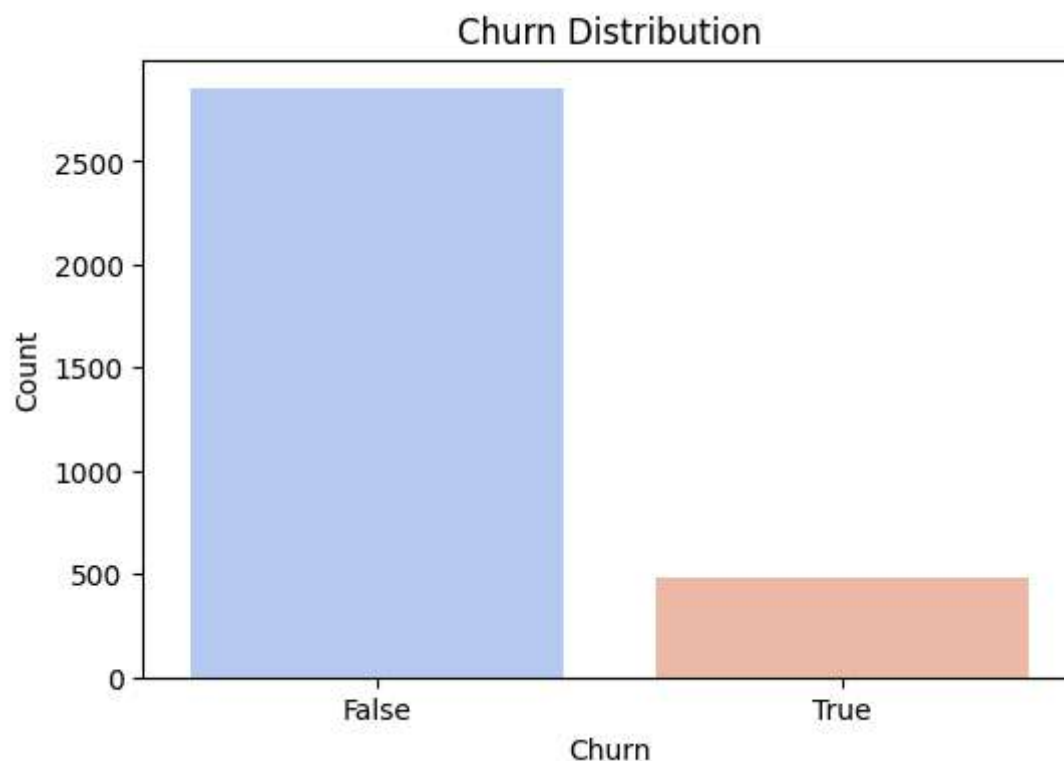
## Dataset Overview

The dataset consists of customer attributes, including:

- **Demographics:** State, area code
- **Service Usage:** Total call minutes, customer service calls
- **Subscription Plans:** International plan, voicemail plan
- **Billing Details:** Total charges, total calls
- **Target Variable:** Churn (1 = Churned, 0 = Retained)

## Exploratory Data Analysis (EDA) & Visualizations

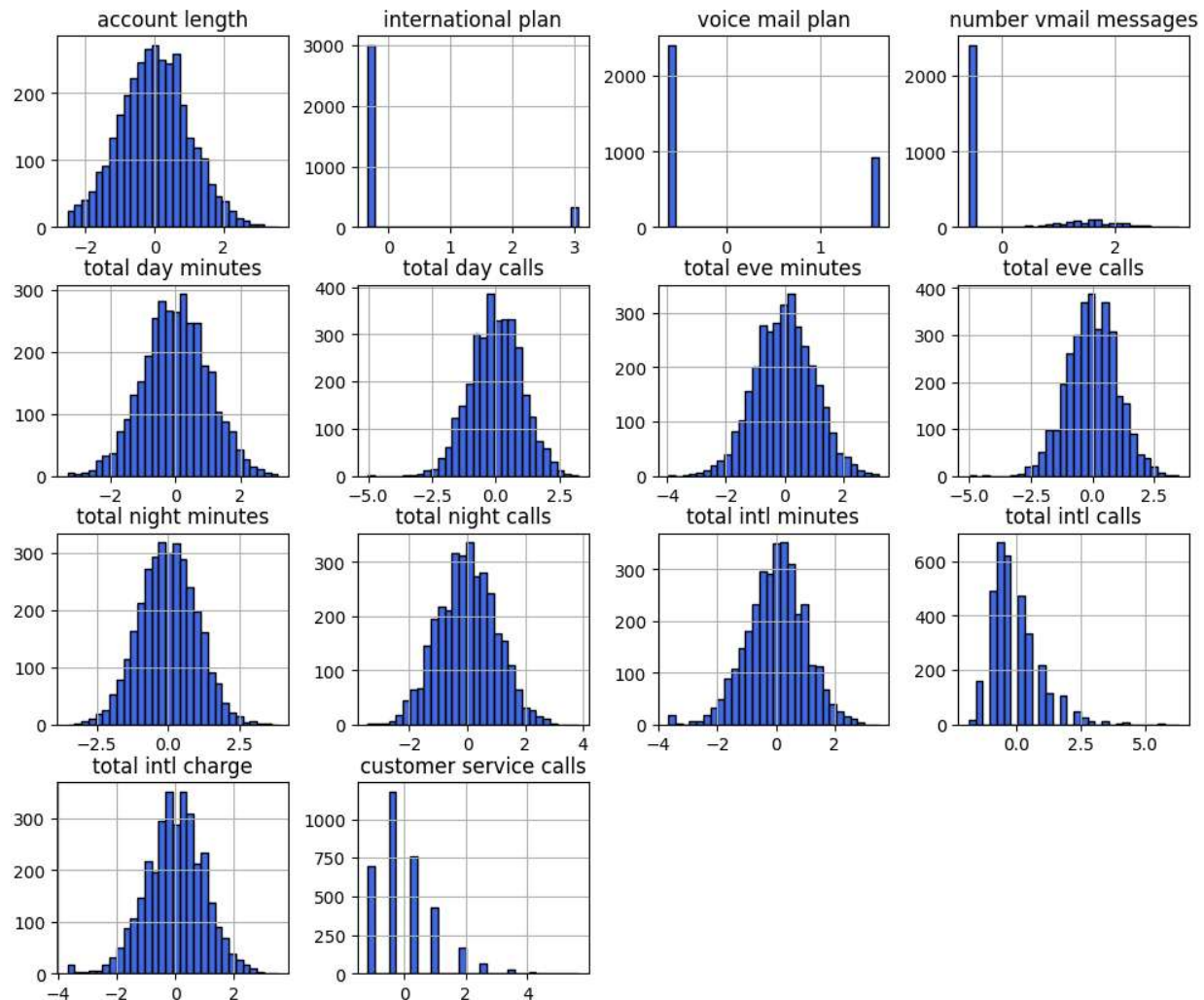
Churn distribution



- The bar graph shows that the dataset is imbalanced, with fewer churned customers (1) than non-churned customers (0). This shows that churn is lower in number, the business might not yet have a severe churn issue but requires proactive measures to prevent escalation.

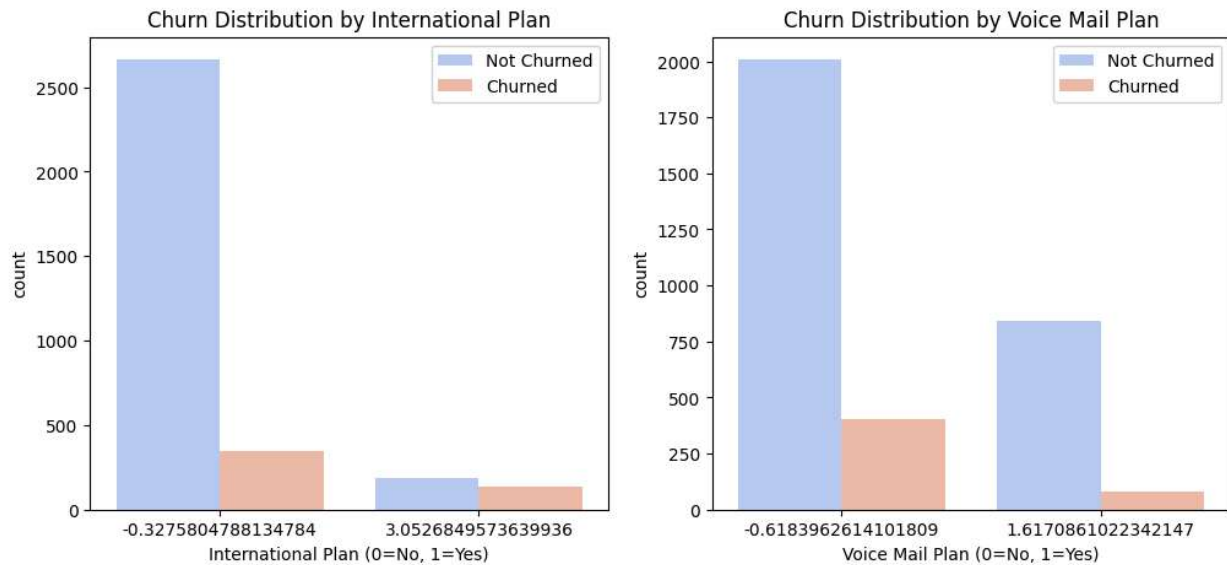
## Distribution of numerical features

Distribution of Numerical Features



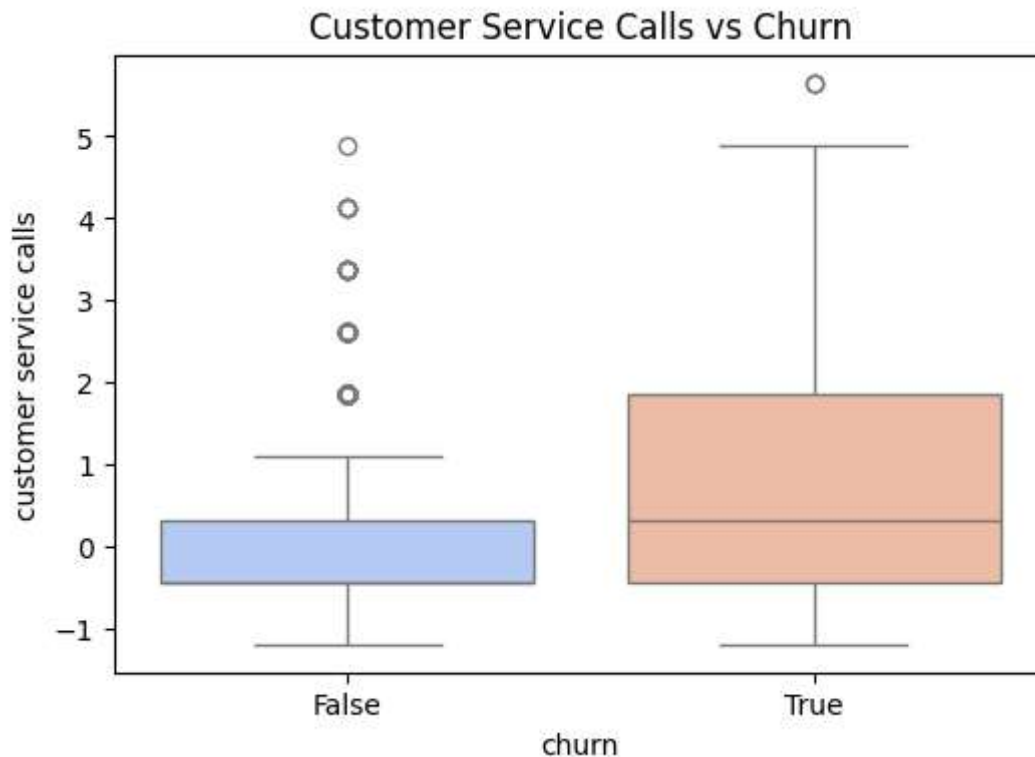
- Features like total day minutes, total eve minutes, and total night minutes are normally distributed. Customer service calls are right-skewed indicating that most customers make fewer calls..

## Churn vs. Categorical Features (International Plan & Voice Mail Plan)



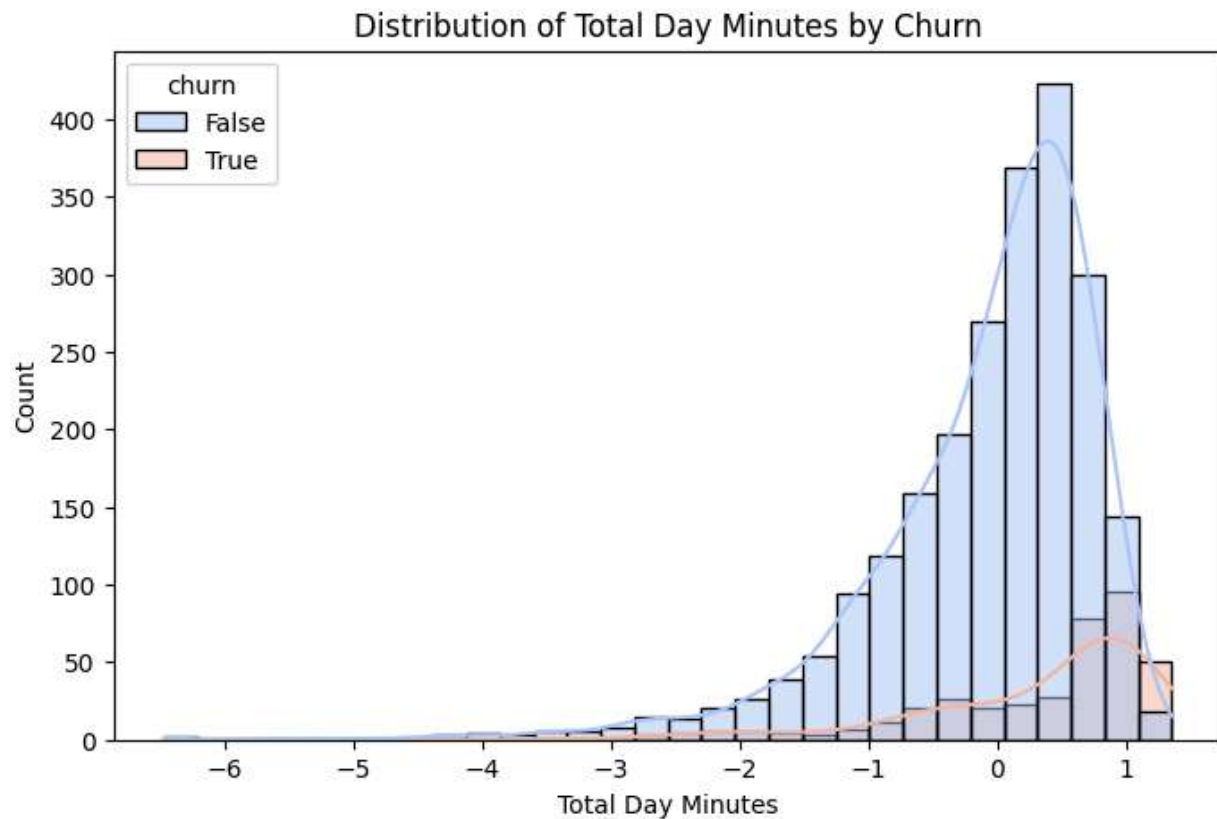
- Customers with an international plan churn more than those without it.
- Voice mail plan does not significantly affect churn..

### Customer service calls vs churn



- The box plot shows that Customers who make more customer service calls tend to churn at a higher rate. Frequent service calls might indicate dissatisfaction with the service.

### Distribution of total day minutes by churn



- The histogram shows a higher density of churned customers in the upper range of total day minutes. Customers who spend more time on calls during the day tend to churn more. This means that high usage customers may be more sensitive to service quality or pricing issues. syriaTel should find ways to customer retention.

## Data Preprocessing

---

### 1. Handled missing values

- converted International Plan & Voicemail Plan to binary values (0 and 1)

### 2. Encoded categorical variables.

- Used one-hot encoding for categorical features like "State" and "Area Code."

### 3. Feature Scaling:

- Applied StandardScaler to normalize numerical columns for better model performance.

## Model Training & Evaluation.

---

**Models Tested:** Logistic Regression, Decision Trees, Random Forest, K-NN, SVM.

**Best Model:** Random Forest, due to its high accuracy and ability to handle imbalanced data.

### Model Performance Metrics:

- **Accuracy:** 95.95%
- **Precision:** 96%
- **Recall:** 76%
- **F1-Score:** 85%

## Key Findings

---

### 1. Customer Service Calls:

- High customer service calls are a strong predictor of churn, indicating dissatisfaction or unresolved issues.

### 2. International Plan:

- Customers without an international plan are more likely to churn, suggesting a need for more attractive international offerings.

### 3. Total Charges:

- Fluctuations in total charges significantly impact churn risk, highlighting the importance of transparent and consistent pricing.

### 4. Usage Patterns:

- High usage customers (e.g., those with high total day minutes) are more likely to churn, possibly due to pricing or service quality concerns.

## Conclusion

---

- Random forest was the best model with 95% accuracy which was the best for predicting accuracy
- By leveraging machine learning to predict customer churn, SyriaTel can take proactive steps to retain customers, reduce revenue loss, and improve overall customer satisfaction. The insights and recommendations provided in this project offer a roadmap for SyriaTel to address churn effectively and enhance its competitive position in the telecommunications market.

# Recommendations & Next Steps for SyriaTel

---

## 1. Improve Customer Support:

- Address the root causes of frequent customer service calls by improving service quality and resolving issues faster.

## 2. Personalized Retention Offers:

- Offer tailored discounts, loyalty rewards, or personalized plans to at-risk customers identified by the model.

## 3. Optimize Pricing Strategies:

- Analyze billing patterns and adjust pricing structures to enhance affordability and customer satisfaction.

## 4. Enhance Service Plans:

- Consider offering flexible international and voicemail plans based on customer usage patterns.

## 5. Proactive Customer Engagement:

- Use the predictive model to identify at-risk customers and engage them with targeted retention campaigns.

## 6. Dashboard Integration:

- Develop an internal dashboard for real-time churn tracking and monitoring of retention strategies.

## References

- <https://moringa.instructure.com/courses> notes
- <https://www.kaggle.com/datasets/becksddef/churn-in-telecoms-dataset>