

Customer Churn Prediction for SyriaTel



Table Of Content

- 1) Overview
- 2) Business Understanding
- 3) Data Understanding
- 4) Modeling
- 5) Evaluation
- 6) Recommendation
- 7) Next steps

Project Overview

- Project objective: Develop a predictive model to identify customers likely to churn for SyriaTel.
- Data source: Utilize a dataset from Kaggle containing customer information relevant to churn prediction.
- Methodology: Employ machine learning algorithms to train models and evaluate their performance.
- Key focus: Rigorous analysis and model selection to ensure alignment with SyriaTel's business objectives.
- Goal: Reduce customer churn and enhance customer satisfaction through proactive retention strategies.

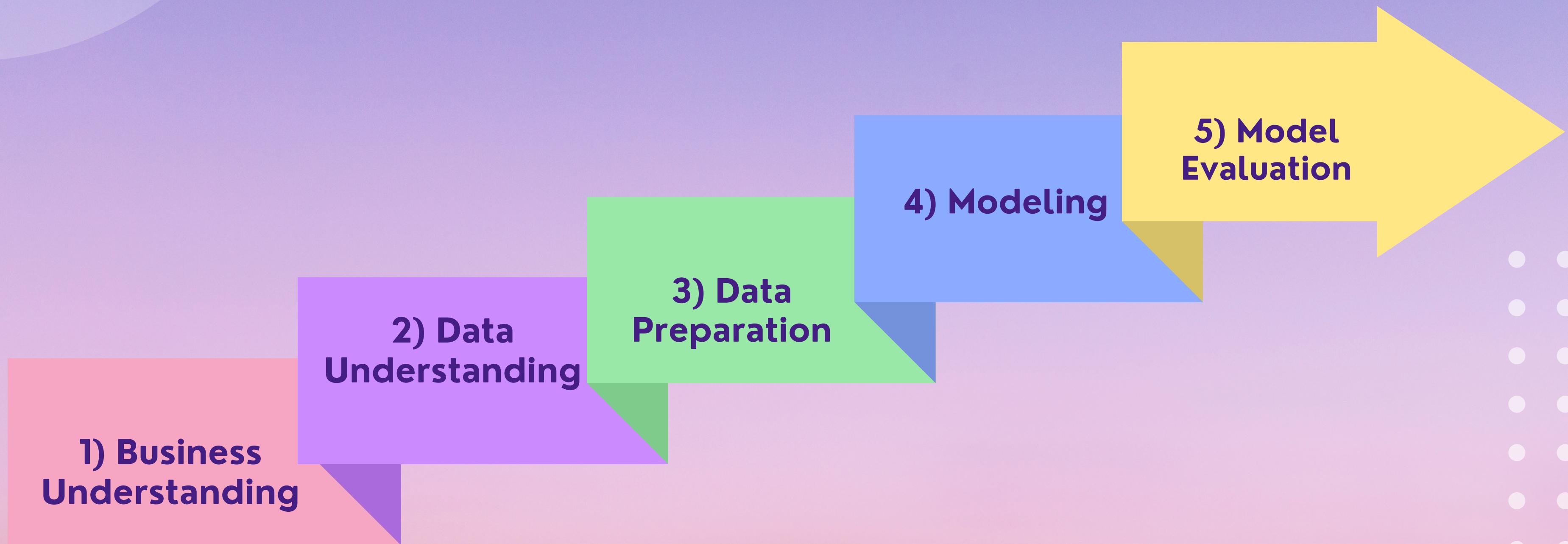
Business Understanding

- Problem Identification: Identify customer churn as a key challenge for SyriaTel, affecting financial performance and market position.
- Objective Definition: Create a churn prediction model to enable proactive retention strategies and minimize revenue loss.
- Business Impact: Enhance revenue, lower acquisition costs, and bolster competitive advantage by improving customer retention.

Data Understanding

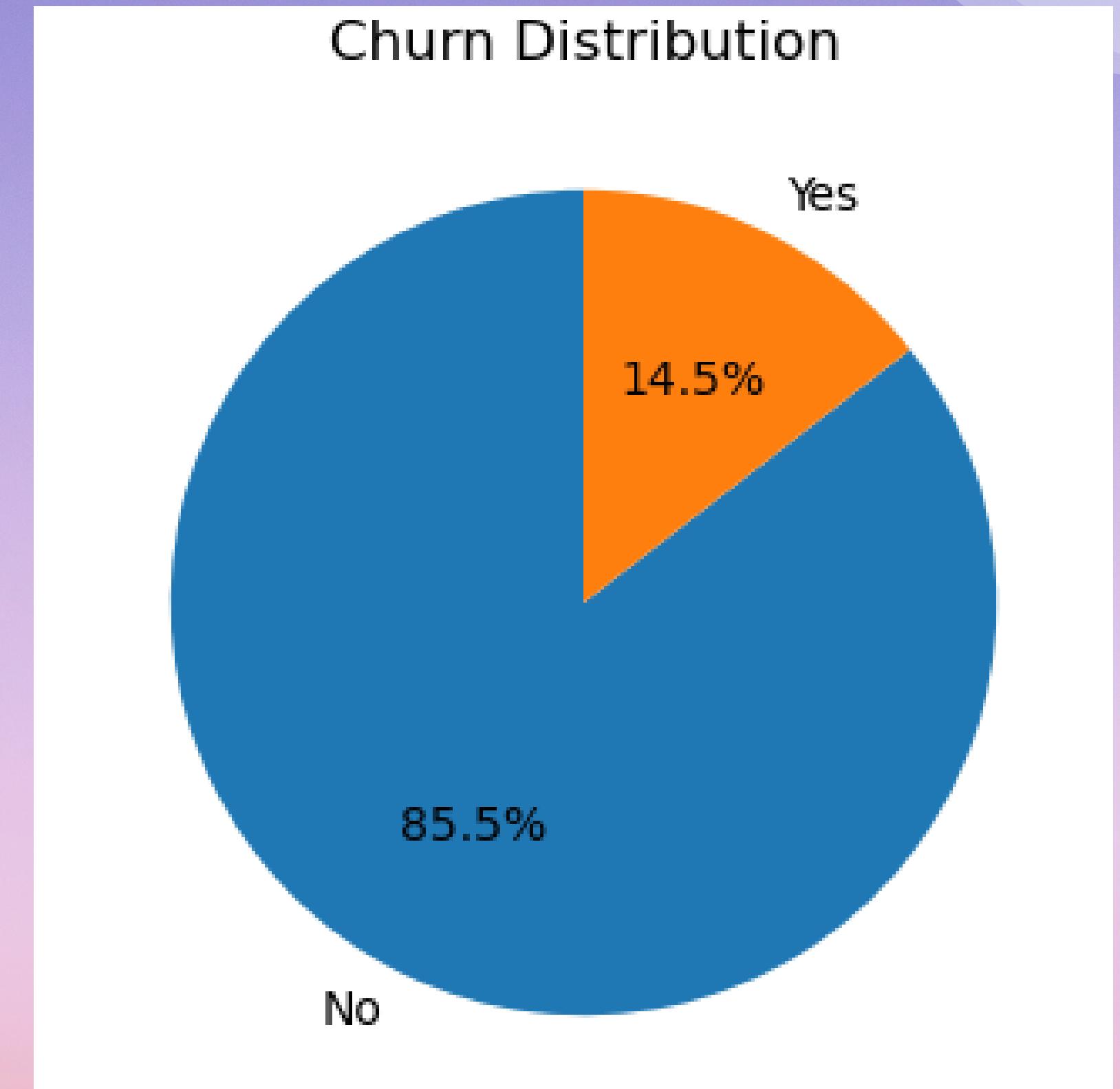
- Dataset Overview: Obtain a dataset from Kaggle containing 3333 rows and 21 columns, representing customer information and churn status.
- Data Cleaning: Address missing values, encode categorical variables, and scale numerical features to prepare the dataset for analysis.
- Exploratory Data Analysis (EDA): Conduct EDA to understand feature distributions and identify potential correlations with customer churn.

Step by Step Process



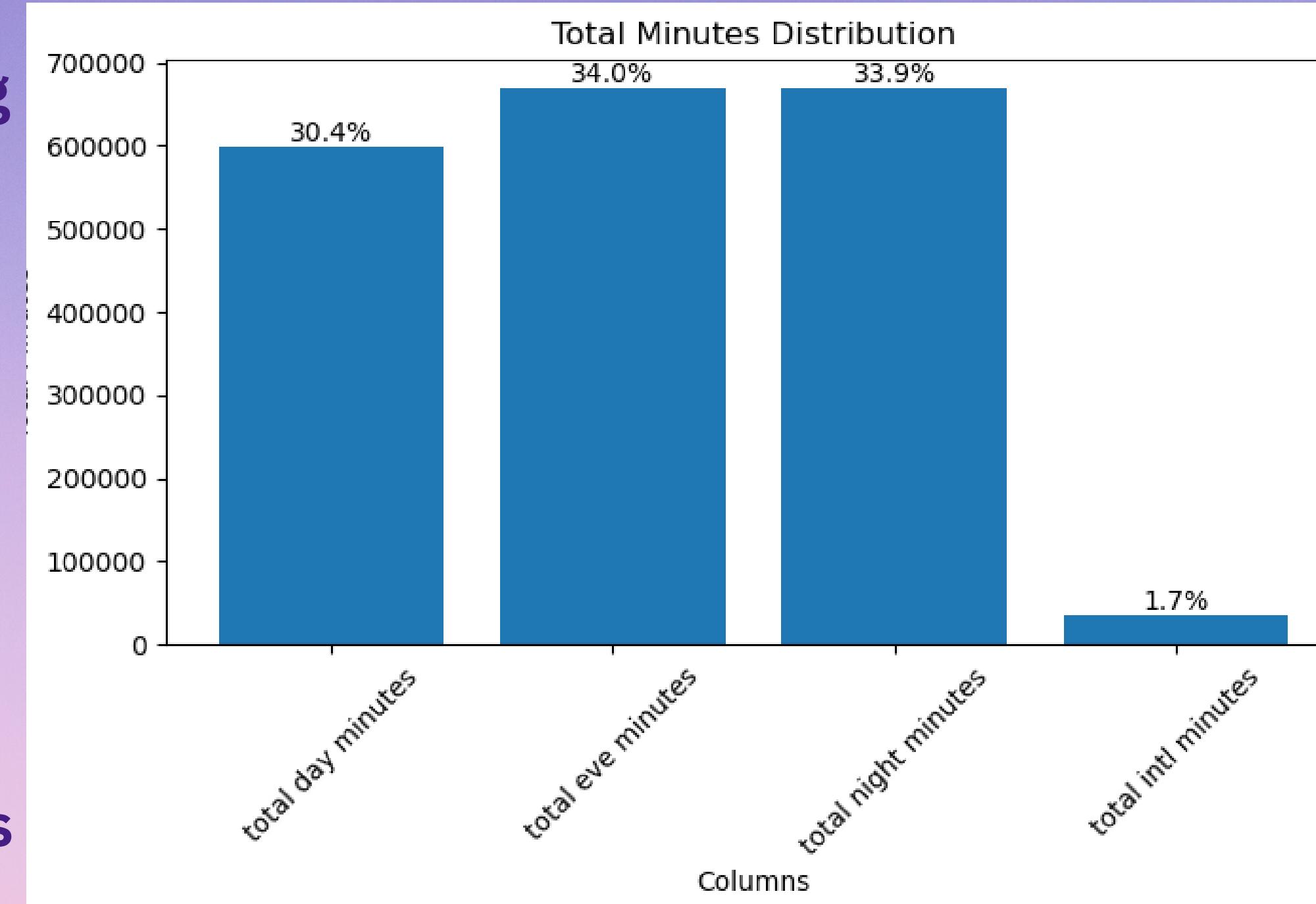
Churn Distribution

**85.5% of the customers
have not churned while
14.5% have churned.**

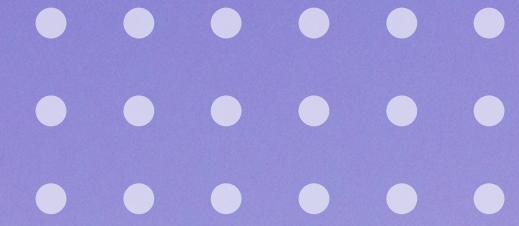


Minutes Distribution

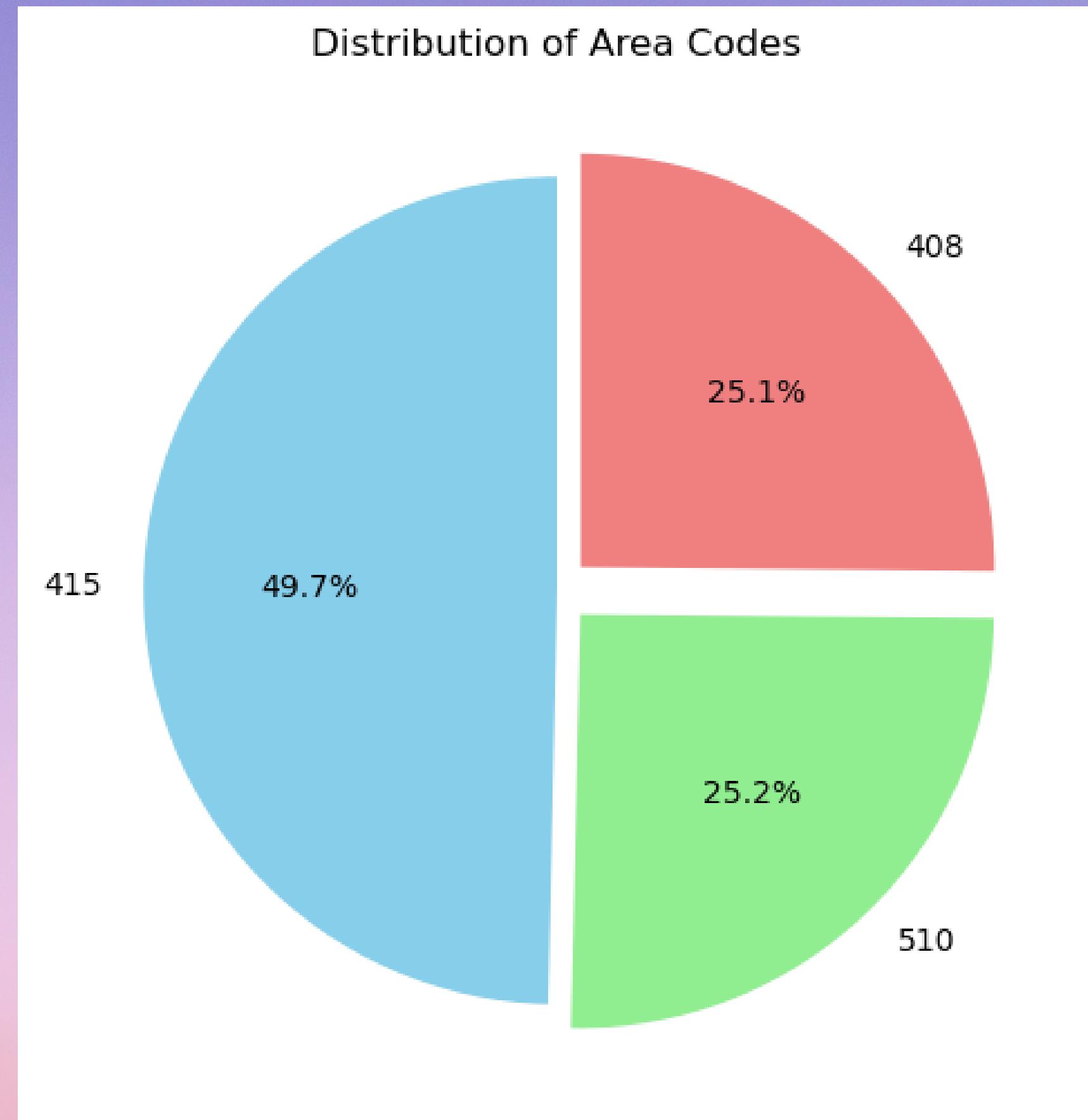
- Evening minutes contribute the most to the overall total minutes, indicating potential usage patterns among customers.
- Understanding the distribution of call minutes can provide valuable insights into customer behavior and preferences.
- SyriaTel can leverage this information to tailor their services and promotions to better meet customer needs and preferences



Churn Distribution By Area Code

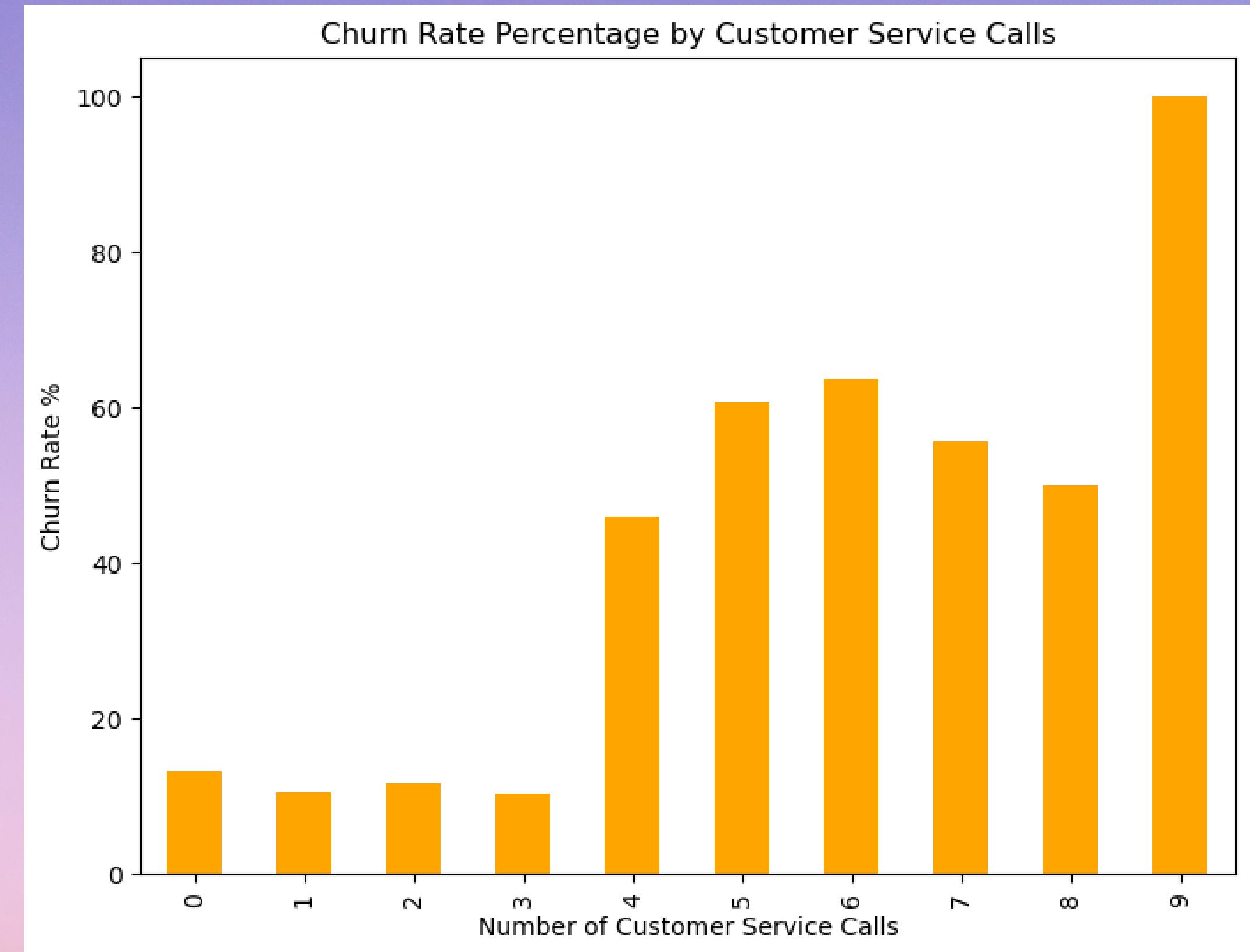


- Area Code 415 exhibits the highest churn rate at 49.7%, highlighting the importance of targeted retention strategies and service quality enhancements in this region



Customer service vs churn

- The more customer service calls made, the higher the chance of churn.
- After the 6th call, churn significantly increases, suggesting a critical point where more calls could lead to more unhappy customers leaving

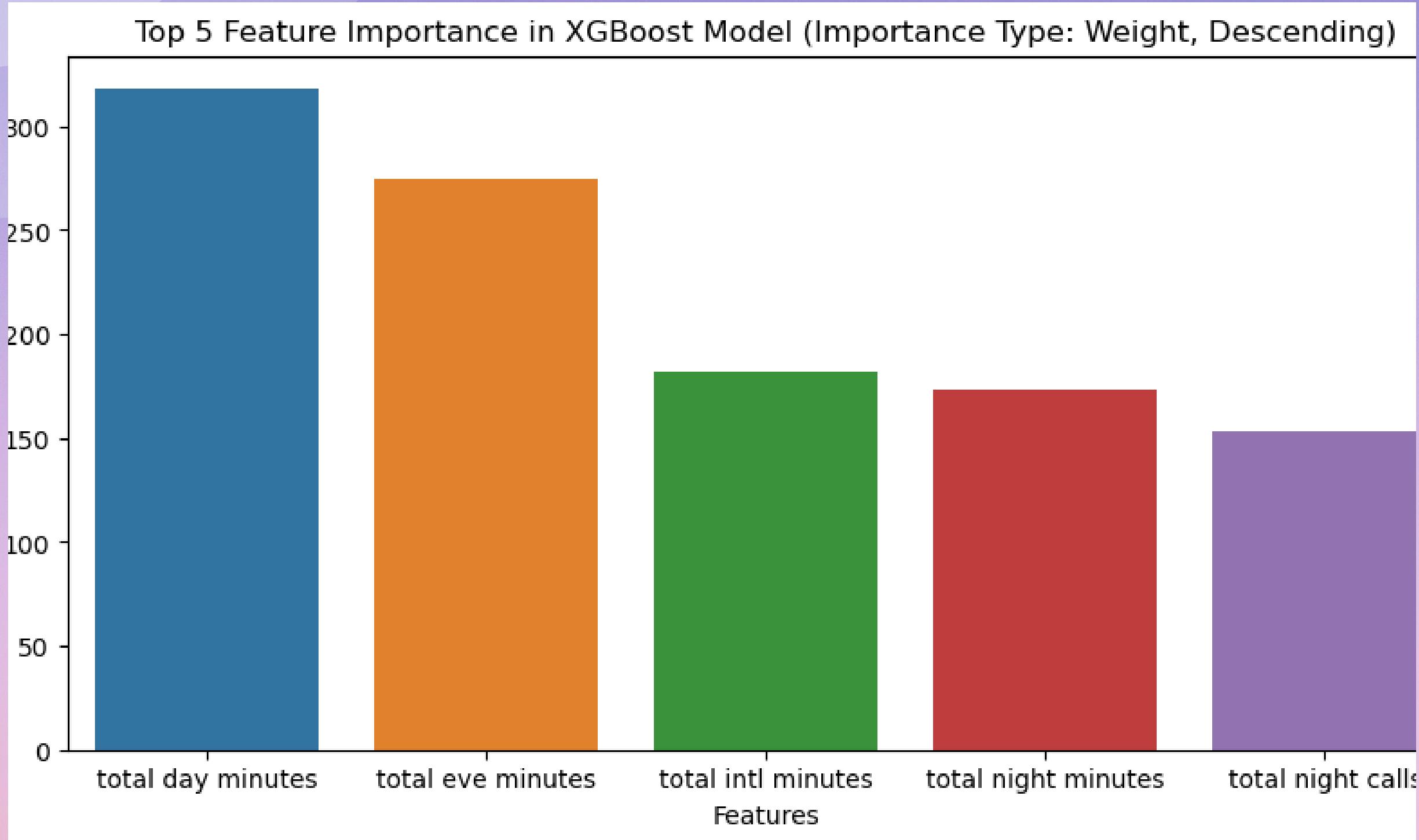


Evaluation

**XGBoost achieves an
impressive accuracy of
95.80%**

Metric	XGBoost
Accuracy	0.958021
Precision	0.950617
Recall	0.762376
F1-score	0.846154
ROC AUC	0.922226

Evaluation Cont



Top 5 Features:
total day minutes
total eve minutes
total intl minutes
total night minutes
total night calls

Recommendation

- **Proactive retention:** Personalized offers, loyalty programs, and targeted marketing.
- **Service improvement:** Enhance quality and address pain points.
- **Effective communication:** Establish reliable channels for feedback and support.
- **Data-driven insights:** Use analytics to understand behavior and inform decisions.

Next Steps

- **Refine Models:** Continuously optimize machine learning models for better predictive performance.
- **Enrich Data:** Explore adding more relevant features to enhance model accuracy.
- **Deploy in Production:** Integrate the final model into operational systems for real-time predictions.
- **Monitor Performance:** Establish robust monitoring to track model effectiveness over time.
- **Data-driven insights:** Use analytics to understand behavior and inform decisions.



Thank you