PREDICTIVE ANALYTICS APPROACH TO UNDERSTANDING CUSTOMER RETENTION

INTRODUCTION

Competition from the telecommunication industry is a great dan ger to business growth. As a result of this, companies in this sector must excel in industry dynamics. One of the best ways to tackle this danger is through increasing customer retention levels.

BUSINESS UNDERSTANDING

- Customer churn in Telecommunication Industry
- Growth and Development
- Predictive models to analyze why customer churn is a major business dynamic in Telecoms
- Audience of the project
- Existence of predictable patterns in Telecoms industry

Stakeholders

- ✓Marketing Department
- ✓Sales Department
- ✓Customer Service Department
- ✓Product Development/Management Team
- ✓Chief executives/Senior Management

OBJECTIVES

- Determine which features or behaviors are most indicative of customer churn.
- Develop a robust predictive model that can accurately identify customers at high risk of churning.
- Provide actionable insights that can help the business develop effective strategies to reduce churn, such as personalized offers, improved customer service, or proactive outreach/marketing.
- Advise the company on resource allocation by targeting retention efforts towards high-risk customers

DATA UNDERSTANDING

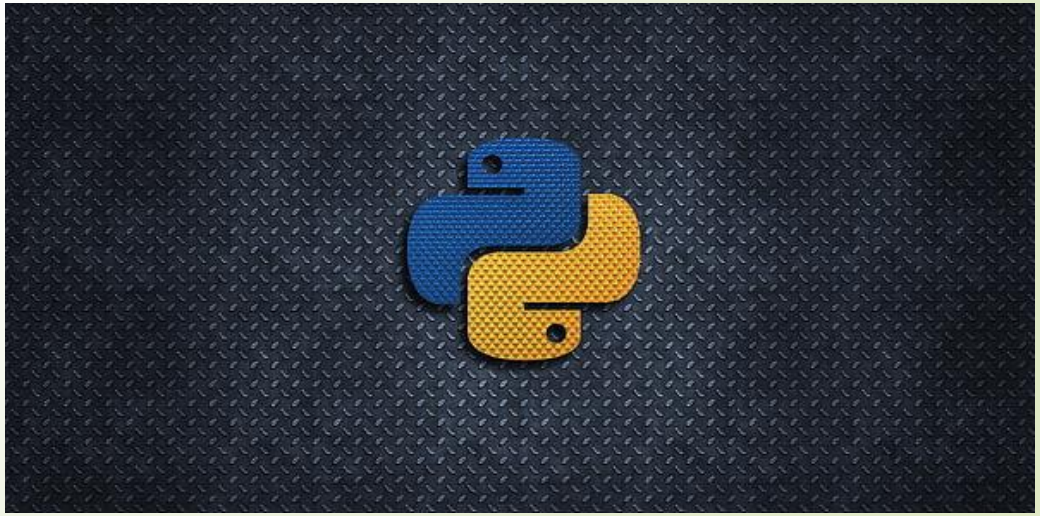
- Source of Data
- Loading Data
- Pandas
- Numpy
- Matplotlib and Seaborn for Visualizations
- Classes in Scikit-Learn for Modeling
- Data Features

DATA PREPARATION

- There are no missing values in this dataset.
- Converted churn (target variable) to numerical (0 or 1).
- Converted 'international plan' and 'voice mail plan' from 'yes'/'no' to numerical (1 or 0).
- Dropped phone number because it is an identifier and not relevant for modeling
- Applied One-Hot Encoding to categorical features: state and area code
- Separate features (X) the predictor variable and (y) the target variable.
- Split the data into training and testing sets.
- Feature Scaling(standardization) to the numerical features to have a mean of 1 and standard deviation of 0

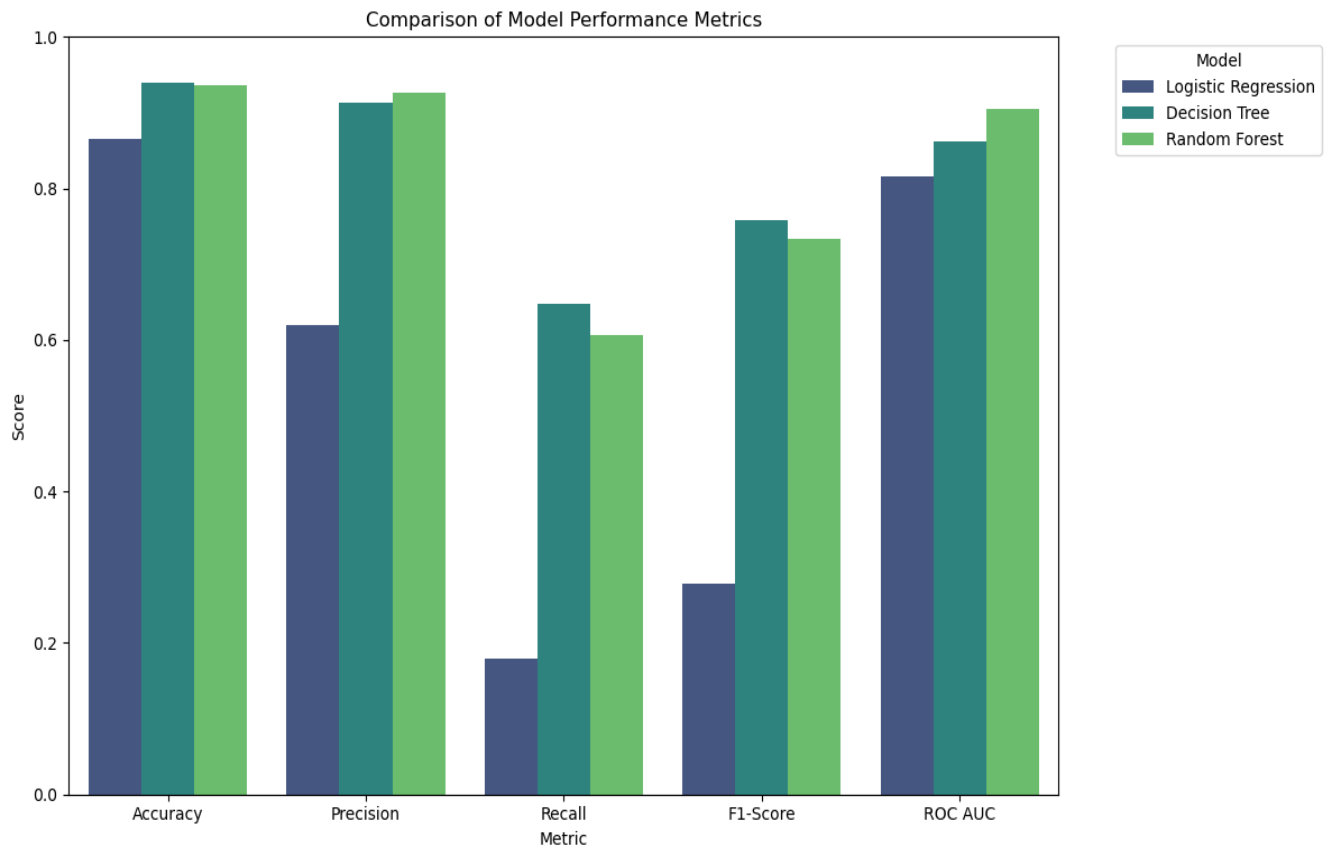
MODELING

- Logistic Regression-Baseline Model
- Decision Tree
- Random Forest



VISUALIZATIONS

Combined Model Performance

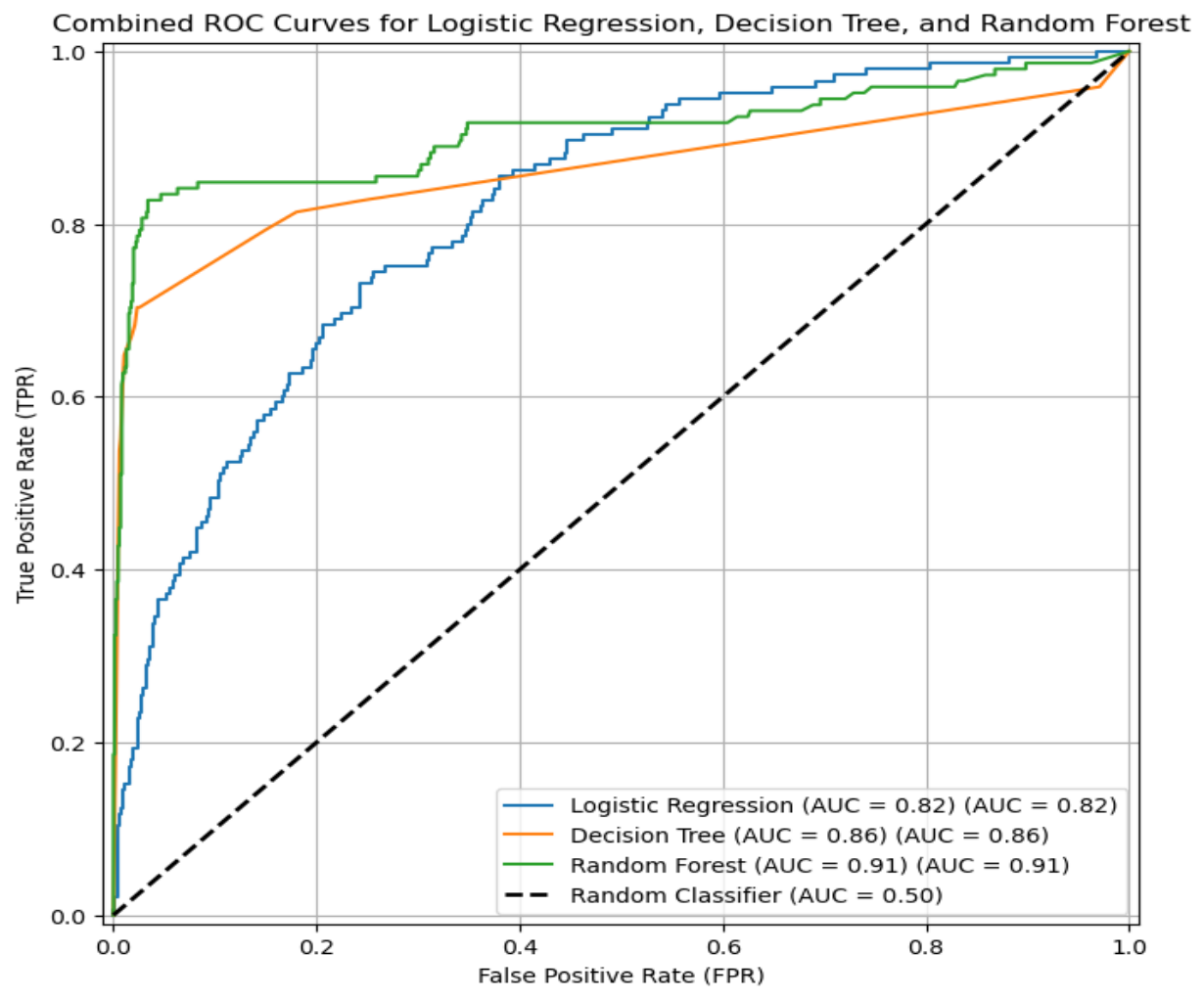


Metrics

- Accuracy
- Precision
- Recall
- F1 Score
- ROC-AUC

ROC – AUC CURVE

Combined ROC-AUC



Metrics

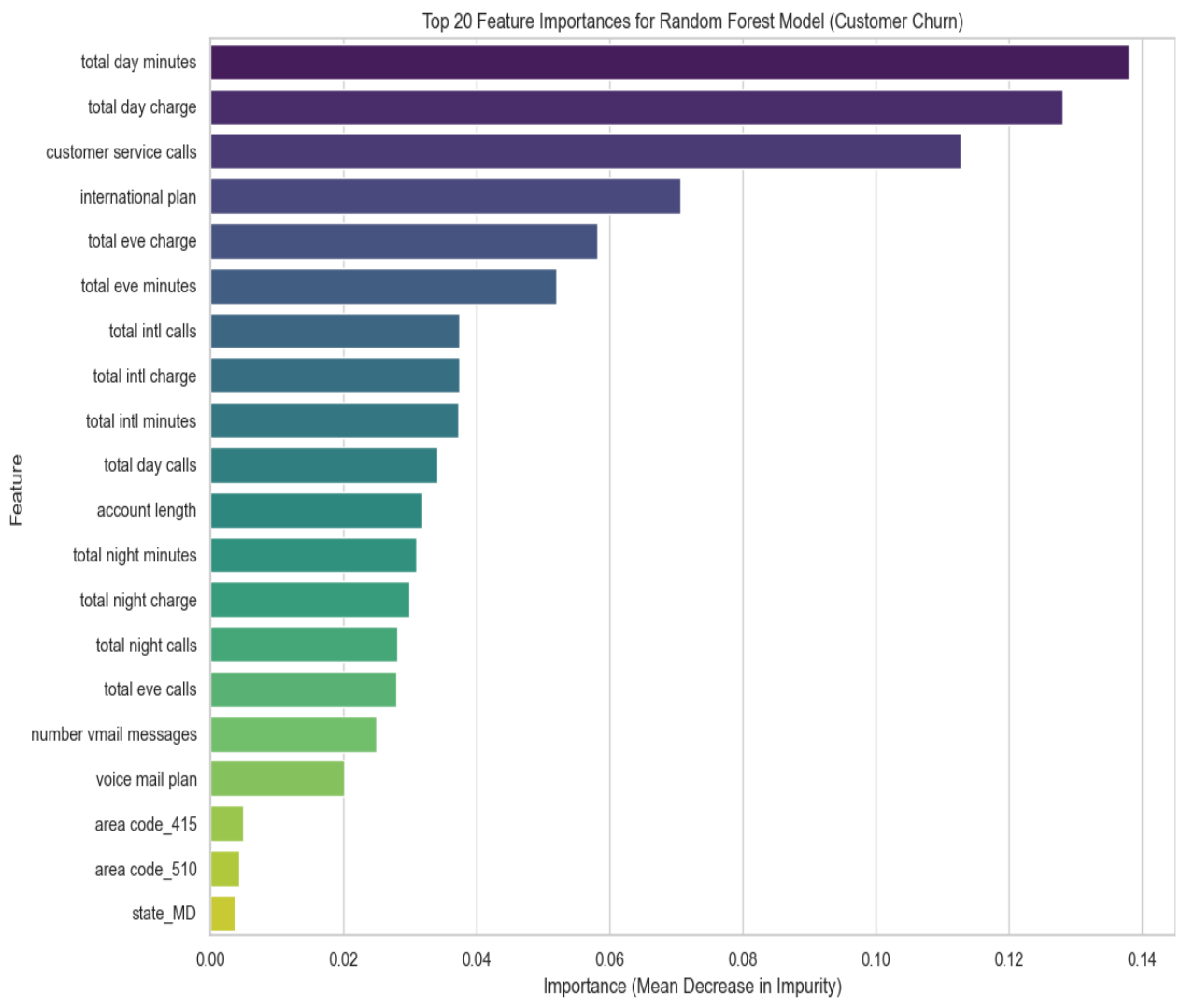
- **Area under the curve**

Random Forest model has very high ability to separate classes

MODEL EVALUATION

- Random Forest has the best overall discriminatory power, meaning it's most effective at distinguishing between churners and non-churners.
- Random Forest has the fewest instances where it incorrectly predicts a non-churner as a churner.
- Random Forest, being an ensemble method, generally provides more stable and robust predictions by averaging out the biases of individual decision trees.
- Random Forest selected for Feature importance

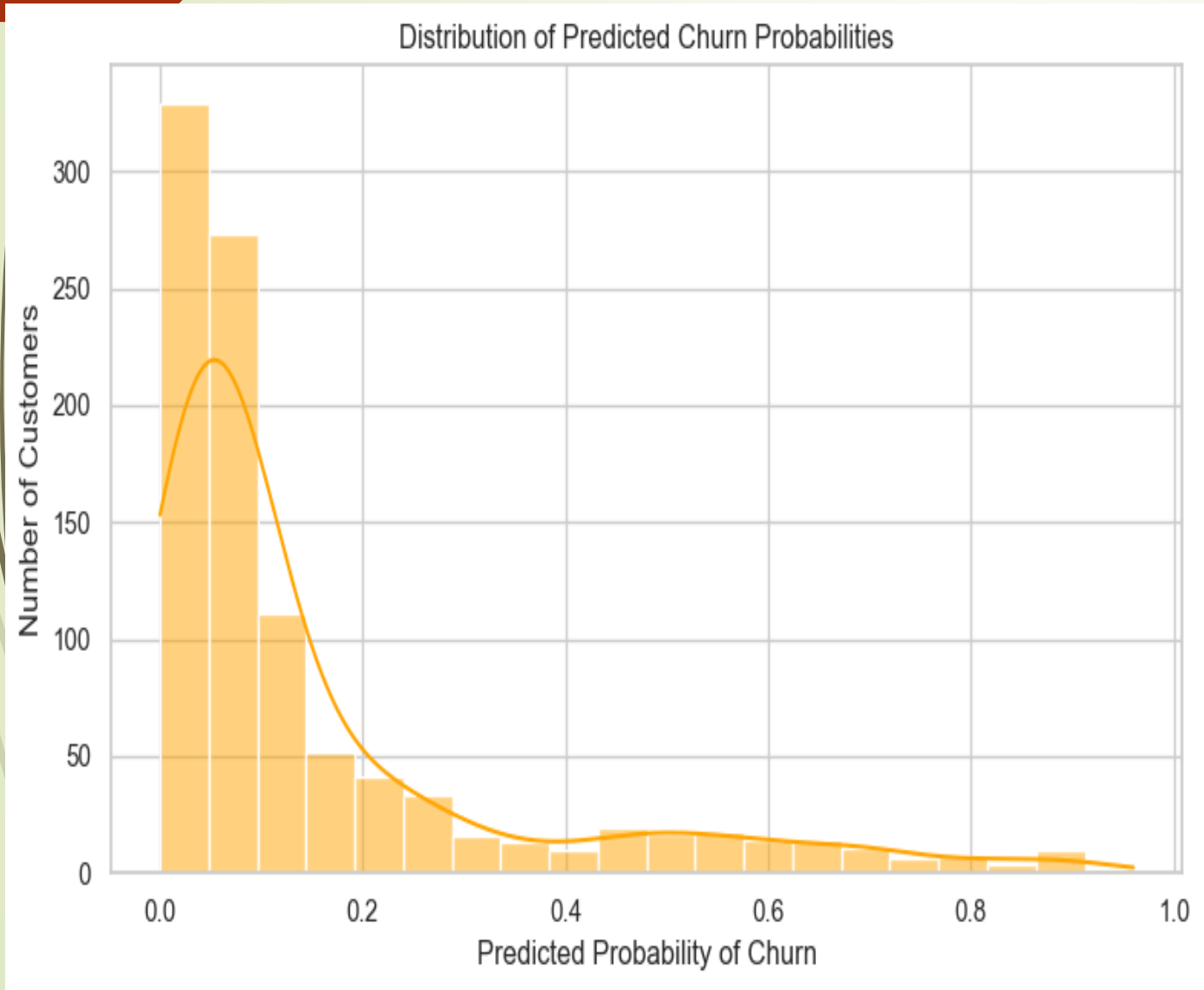
Feature importance



Important Features

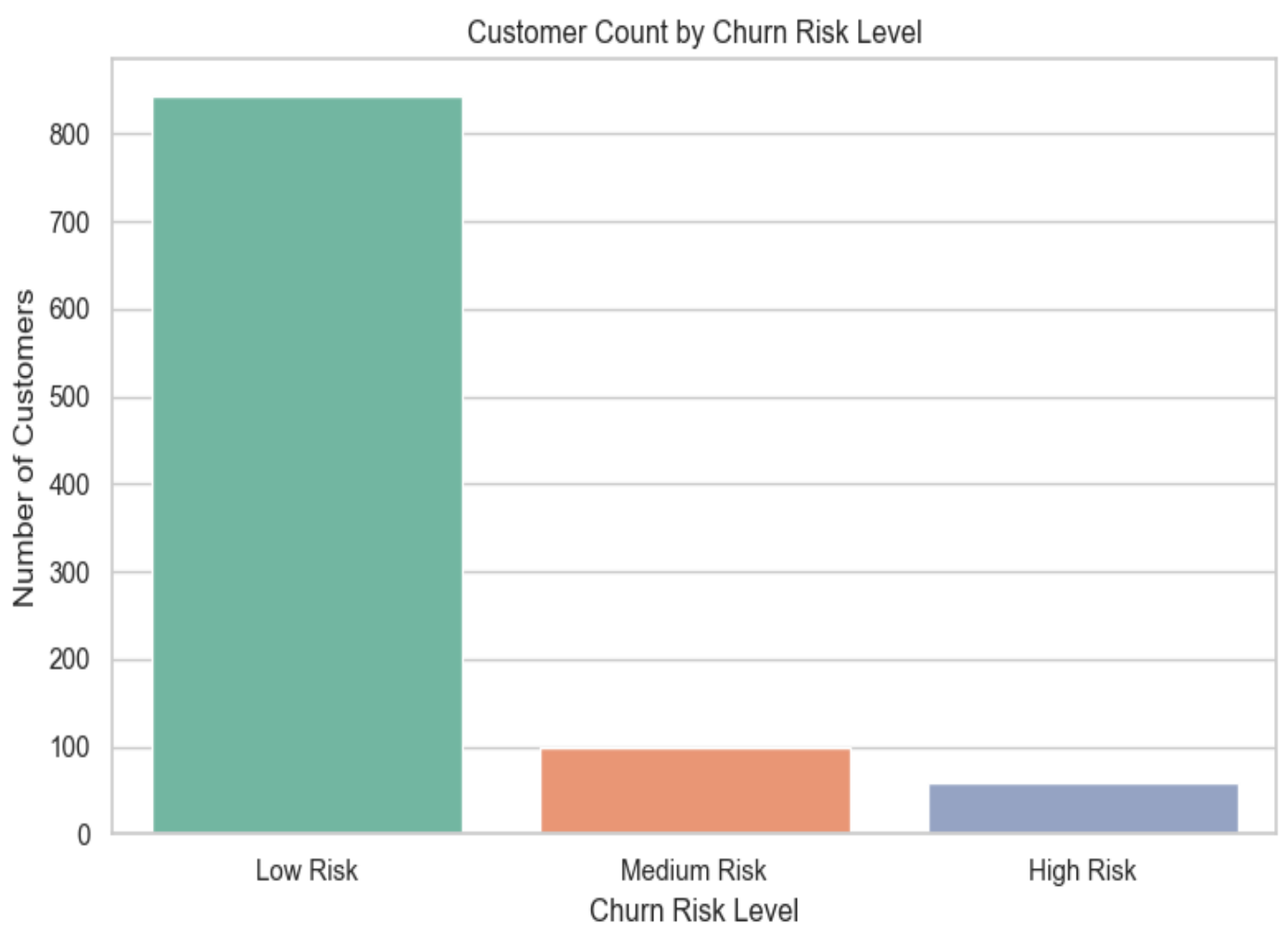
- Total day charge and total day minutes
- The number of customer service calls
- International plans and usage (minutes, calls, charge)
- Evening usage and charges

Churn Probability Distribution



- Most customers have a low predicted probability of churn (between 0.0 and 0.2
- churn is relatively rare in the dataset.
- A smaller portion of customers fall within the medium to high churn probability range (above 0.4)

Customer count vs Risk Level



Low Risk (0.00–0.29)

- The majority of customers fall into this category

Medium Risk (0.30–0.59)

- A smaller segment of customers is at moderate risk

High Risk (0.60–1.00)

- Critical group with a high likelihood of churning

RECOMMENDATIONS

- Total day minutes, customer service calls, international plan, total evening charges are the most important features in determining the customer churn.
- Random Forest is recommended as the most robust predictive model that accurately identified customers at risk but need to be retrained periodically using latest data to mention prediction accuracy.
- Assign customer success managers or special support channels to high-risk accounts Use churn probabilities to create automated alerts for sales or customer care teams when a customer's risk rises.
- Develop a marketing strategy by offering exclusive benefits, discounts, or loyalty points low-risk customers.
- Track engagement signals (e.g., reduced usage, late payments) to detect early signs of churn.
- Launch urgent, personalized outreach campaigns
- More resources to be allocated to customers categorized as high risk, and those in medium risk level because this are potential churners.

THANK YOU

Q & A