MORINGA

Discover · Grow · Transform

DSF-PT05

PHASE 3 PROJECT

GROUP 13

THE TEAM

- MIRIAM
- MOSES
- STEPHANIE
- PHILIP
- CHEPKEMOI
- MARK

UNDERSTANDING CUSTOMER CHURN, A CASE OF SYRIATEL COMMUNICATIONS

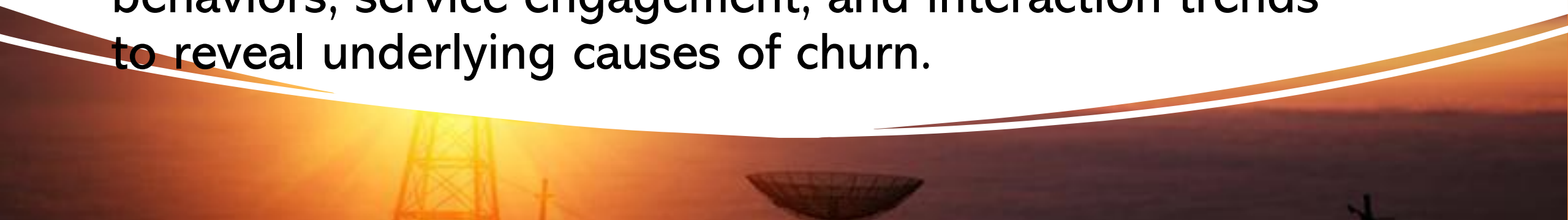
A Deep Dive into Telecom Customer Behavior

Project Overview

Purpose: To tackle the challenge of rising customer churn by leveraging data-driven insights.

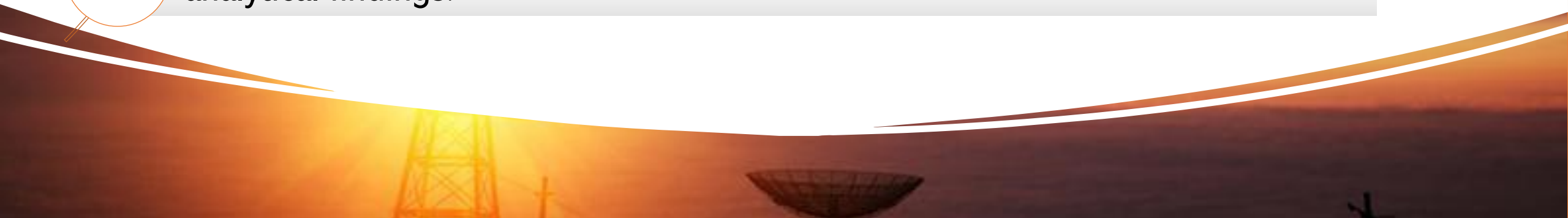
Goal: Identify and mitigate key factors leading to customer departure.

Method: Conduct a comprehensive analysis of customer behaviors, service engagement, and interaction trends to reveal underlying causes of churn.



Objectives

- To utilize clustering techniques to segment SyriaTel's customer base into distinct groups based on usage patterns, demographics, and other relevant factors.
- To assess the impact of various service features on customer retention using predictive modeling techniques.
- To validate the predictive model's performance using appropriate evaluation metrics and refine it to improve accuracy and generalizability.
- To evaluate the long-term effectiveness of implemented retention strategies through monitoring and analyzing churn rates over time.
- To formulate and recommend bespoke retention strategies based on the analytical findings.



Data Pre-processing & EDA

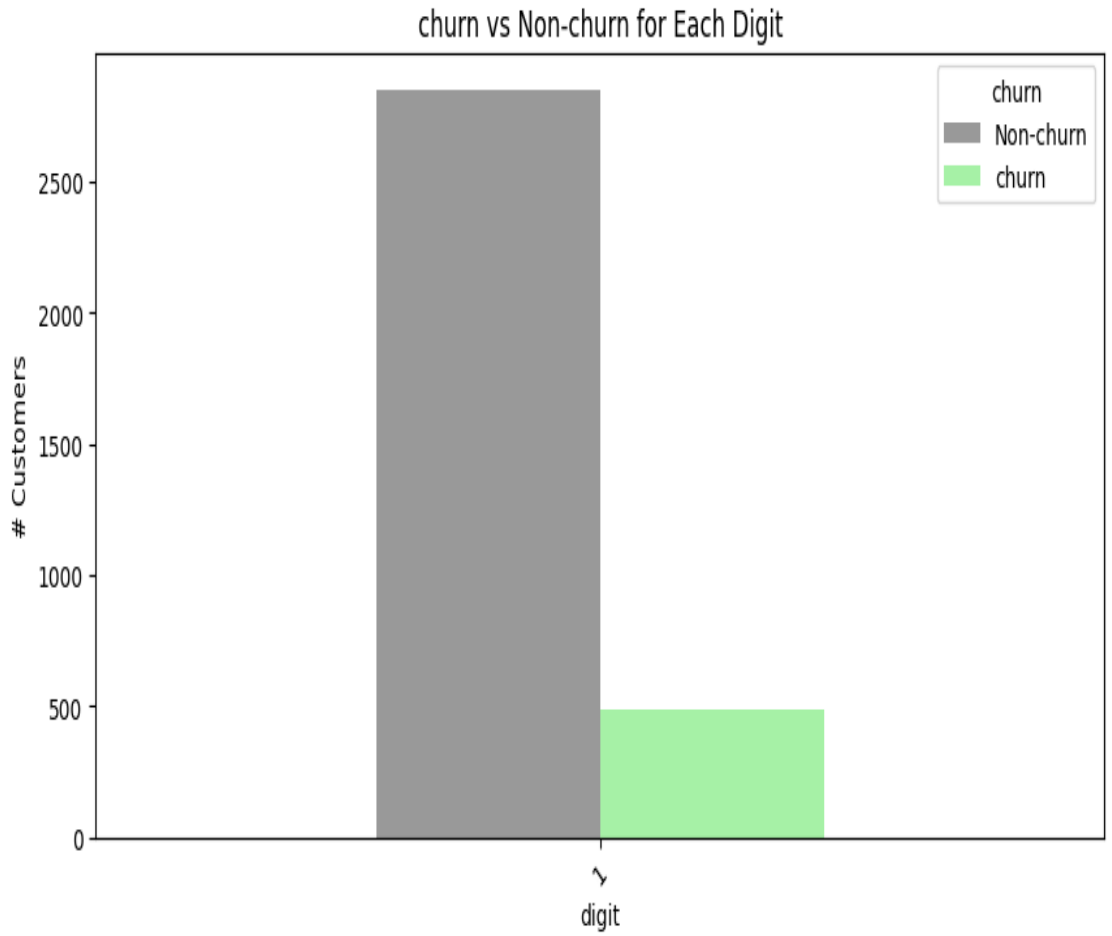
An analysis to extract meaningful insights from the data and identify the best features to be used for modeling

Data preprocessing

- Assumptions about data shape
- Missing values
- Data types
- Categorical variables
- Outliers or errors
- Feature Engineering/Creating

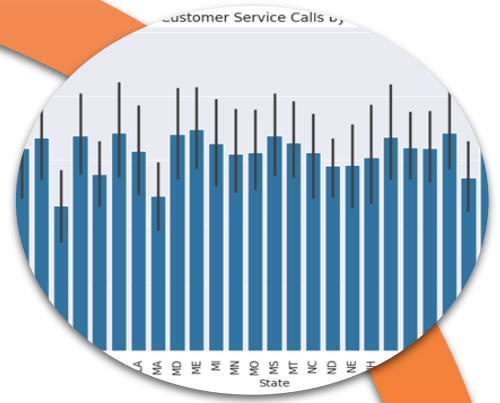
EDA

- Statistical Summary
- EDA Univariate Analysis
- EDA Bivariate Analysis
- EDA Multivariate Analysis
- Conclusion

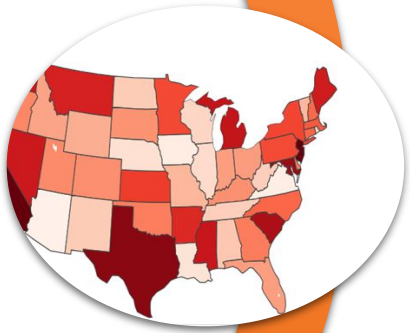


Pre- processing visualizations

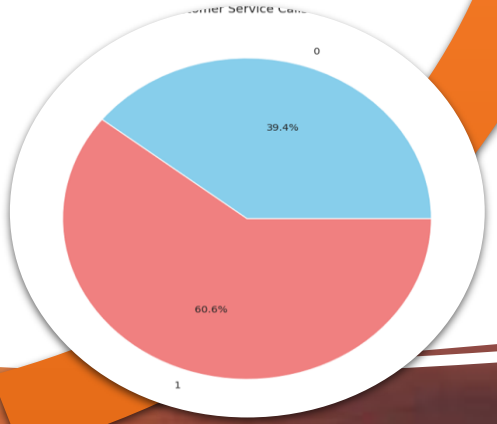
Visualizations are critical to providing insights into the relationships between different variables and the target variable (churn).



Customer Service calls by state



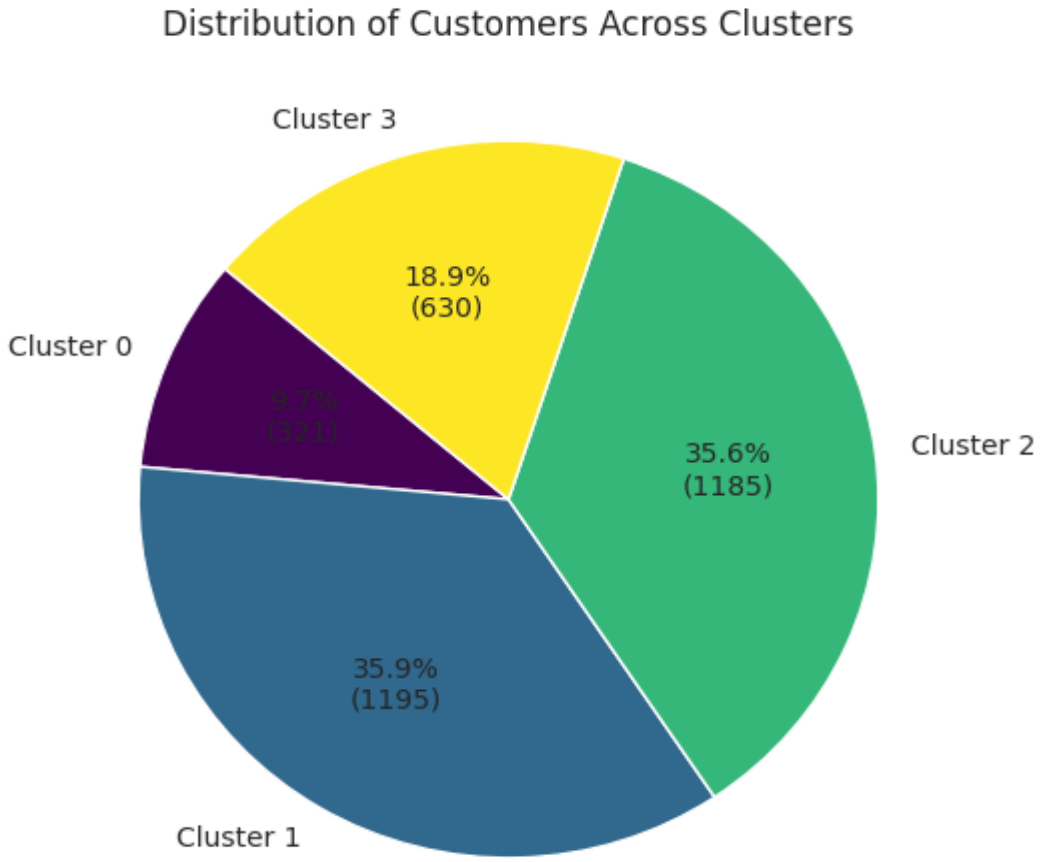
Customer Churn by State

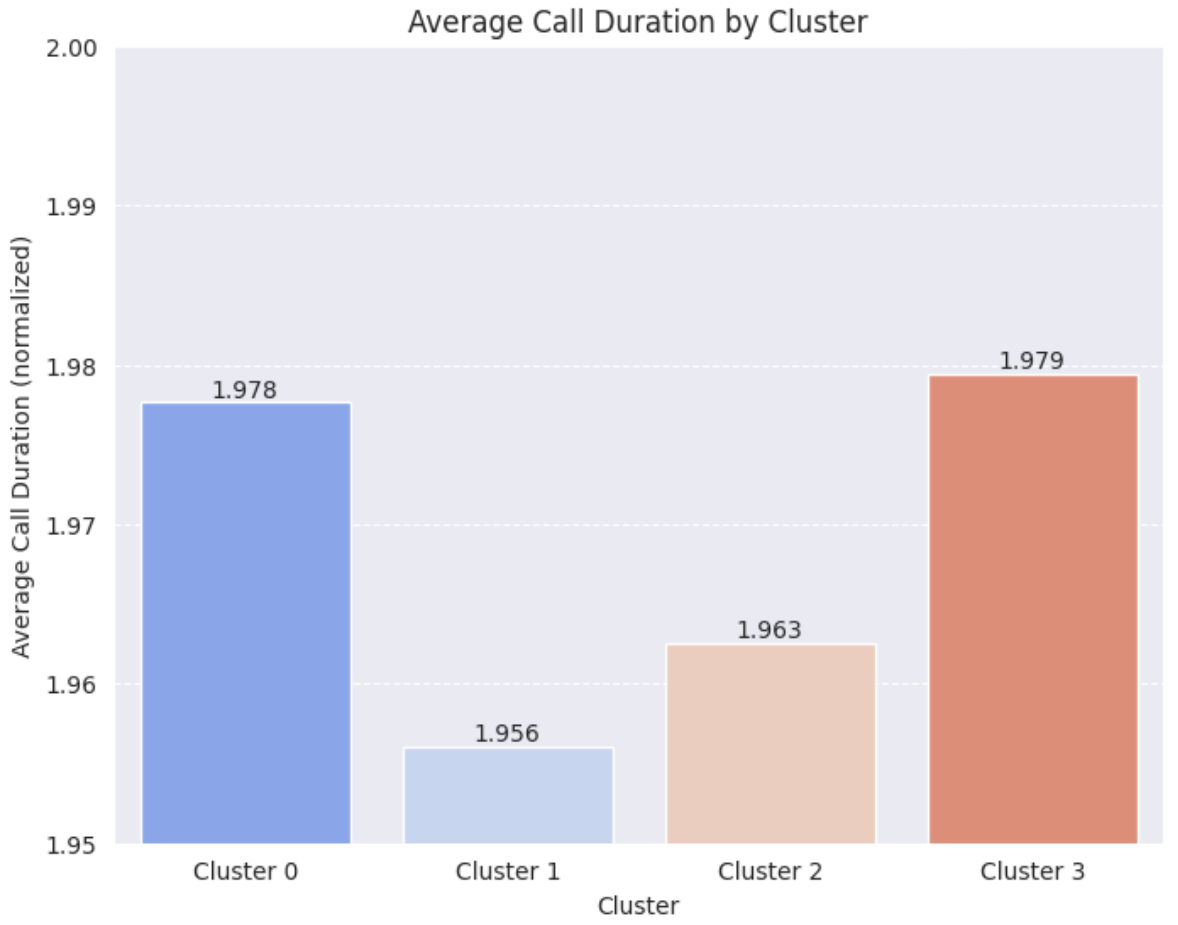


Mean customer service calls by churn

Segmentation and Clustering

Use Elbow Method to determine the most appropriate number of clusters for K-means.
Features selected for Clustering: Total Usage Minutes, Total Calls, Customer Service Calls, International Plan, and High International Usage with some of the results below.

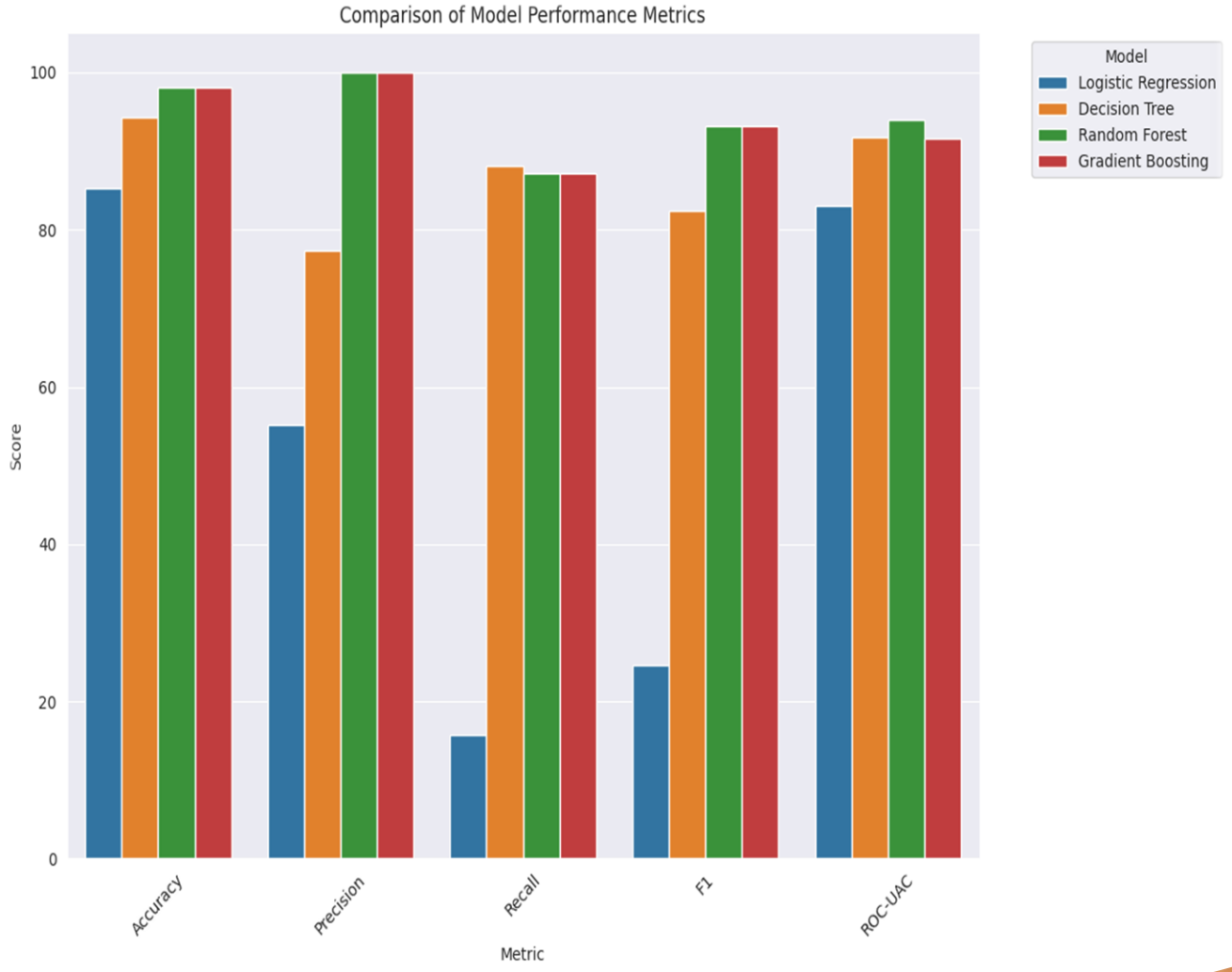




Modelling

Taking an iterative approach to the classifier models and using 3 models to analyze the accuracy, precision, recall, and F1 scores, with ROC/AUC Score and providing a classification report for;

a. Logistic Regression
b. Decision Tree Classifier
c. Random Forest Classifier
d. Gradient Boosting Classifier



Model Performance Overview

Gradient Boosting Classifier

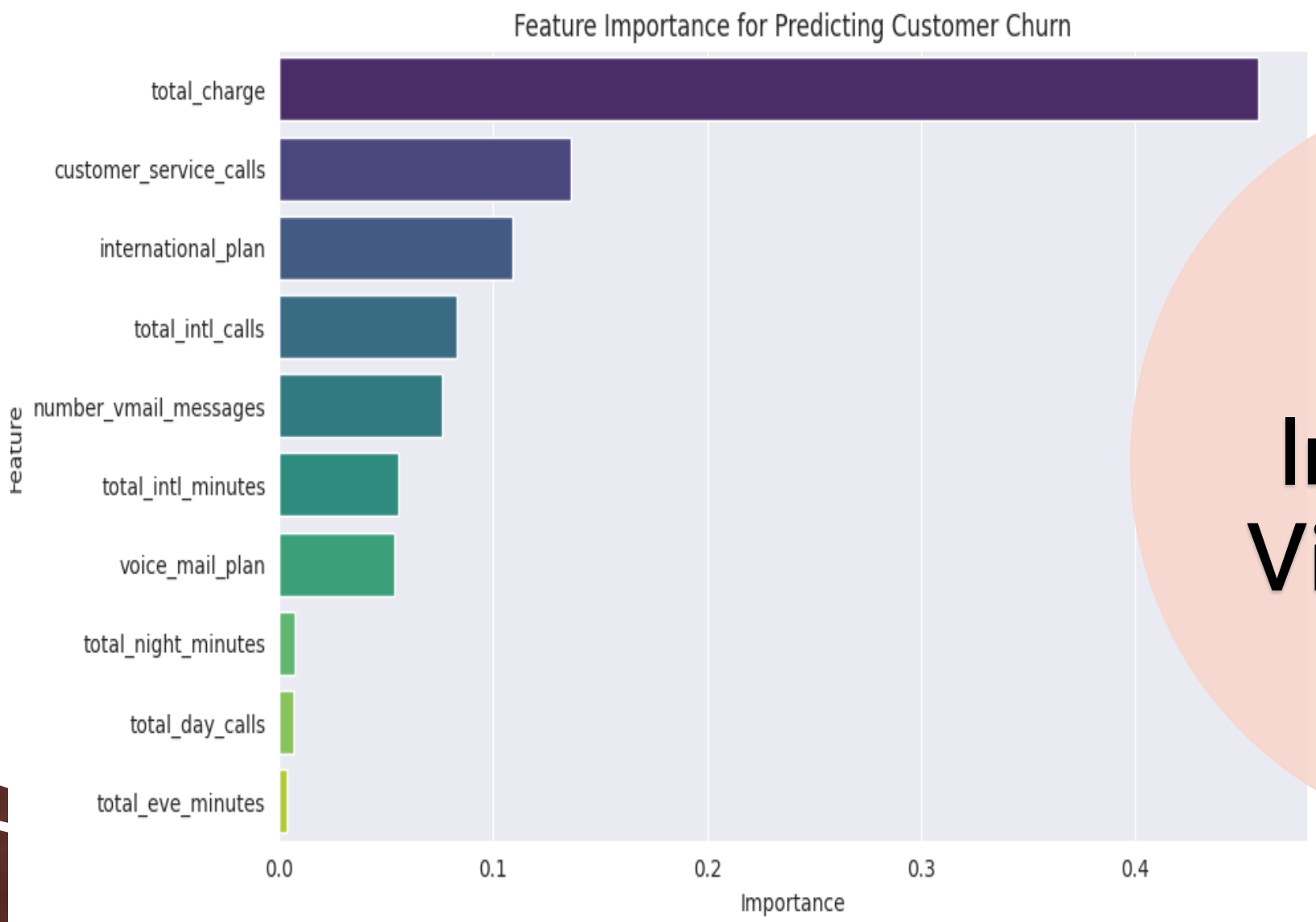
- Highest accuracy for churn prediction.
- Slightly lower ROC-AUC score vs. Random Forest suggests limitations in prediction confidence across thresholds.

Random Forest Classifier

- A strong balance of accuracy and class distinction capabilities (high ROC-AUC score).
- Robust choice for churn prediction, offering reliable performance.

Logistic Regression

- Offers simplicity and interpretability.
- Valuable for scenarios requiring insight into the influence of individual features, despite lower performance metrics



Feature Importance Visualization

Conclusions

Churn Analysis

•**Churn Rate:** At 14.5%, it highlights a notable customer departure rate.
•**Influential Factors:**International plans and the frequency of customer service calls exhibit a moderate positive correlation with churn, significantly impacting customer retention efforts.

Customer Behavior Insights

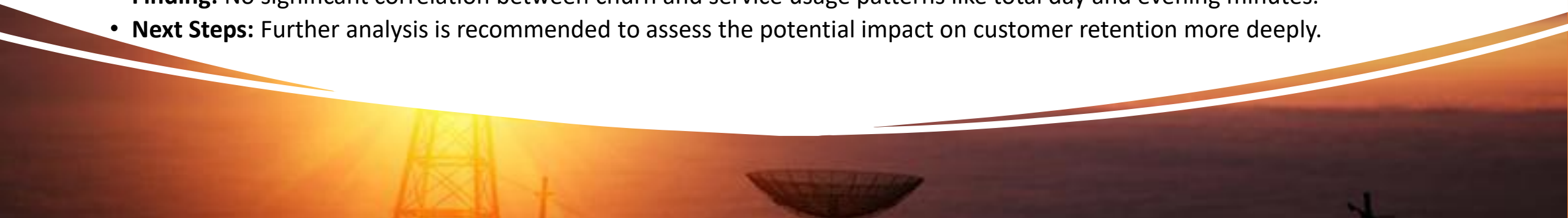
- **International Plan Impact:** Customers with international plans show higher churn rates, signaling a need for deeper analysis on dissatisfaction causes.
- **Customer Service Calls Correlation:** A weak positive correlation with churn implies that increased customer service interaction might not directly enhance retention.

Geographical Patterns

- **Observation:** Churn rates vary across different states, indicating potential geographical influences on customer behavior.
- **Action Item:** Further exploration needed to understand regional differences and tailor retention strategies accordingly.

Service Usage Analysis

- **Finding:** No significant correlation between churn and service usage patterns like total day and evening minutes.
- **Next Steps:** Further analysis is recommended to assess the potential impact on customer retention more deeply.



Recommendations

- Targeted Retention Strategies
- Enhanced Customer Support
- Geographically Targeted Campaigns
- Continuous Monitoring and Analysis
- Data-driven Decision Making
- Collaboration Across Departments

Model Deployment App