PITCH DECK

SUSAN MWORIA

INTRODUCTION

Context:

SyriaTel, like most telecom companies, faces challenges with customer retention. A high churn rate can have significant financial implications.

Objective:

The goal of this analysis is to understand customer churn and develop strategies to reduce it, enhancing customer retention and minimizing revenue loss.

PROBLEM

HIGH CUSTOMER CHURN RATE

SyriaTel is facing a high customer churn rate, meaning a significant percentage of customers are leaving their services. This leads to a loss in revenue and increases the cost of acquiring new customers.

LACK OF PREDICTIVE INSIGHT INTO CUSTOMER BEHAVIOR

SyriaTel currently lacks accurate predictive models to forecast which customers are at risk of churning. Without this insight, the company cannot take timely actions to prevent churn.

INEFFICIENT USE OF MARKETING AND RETENTION RESOURCES

SyriaTel may be spending its marketing and retention budget inefficiently by targeting customers who are unlikely to churn, instead of focusing resources on high-risk customers who require immediate attention.

INABILITY TO OFFER PERSONALIZED CUSTOMER SOLUTIONS

SyriaTel lacks the ability to offer personalized retention strategies tailored to individual customer needs, leading to ineffective attempts at keeping customers.

HOW DO WE SOLVE THE PROBLEM:

1

DATA PREPARATION AND EXPLORATION

Action: Gathered and cleaned customer data, handling missing values, outliers, and encoding categorical variables.

Outcome: Identified key features influencing churn

2

MODEL DEVELOPMENT AND EVALUATION

Action: Tested various machine learning models (e.g., Random Forest, XGBoost) and optimized them using hyperparameter tuning and SMOTE for class imbalance.

Outcome: Selected the best-performing model, balancing accuracy, recall, and precision to predict churn.

3

RECOMMENDATIONS AND SOLUTIONS

Action: Based on model insights, recommended targeted actions to reduce churn.

Outcome: Provided actionable steps to enhance customer retention and minimize churn-related revenue loss.

KEY FEATURES AFFECTING CHURN

TOTAL CHARGES

Customers that had more calls, talk time in minutes had highest charges and are at higher risk of churn

INTERNATIONAL PLAN AND VOICE MAIL PLAN

Customers subscribed to the international call plan and Customers subscribed for voice mail messages

CUSTOMER SERVICE CALLS

Customers that frequently called to the call centre for support

MODEL PERFORMANCE

Performance metrics for best performing models:

Model	Accuracy	Precision	Recall
Tuned Random Forest:	92.2%	93%	52%
Tuned XGBoost	92.9%	97%	55%

NB:
Accuracy: Overall correctness of model
Precision: Positive predictions that are correct
Recall: Proportion of actual positives identified correctly

SOLUTION

IMPROVE CUSTOMER RETENTION PROGRAMS:

Focus on high-risk customers identified by the models. Offer targeted promotions, loyalty programs, and personalized offers to keep them engaged.

PROACTIVE CUSTOMER SUPPORT:

Provide dedicated support for customers identified as 'at-risk' (e.g., high number of customer service calls.)

Introduce early intervention strategies when customers exhibit signs of dissatisfaction.

PRODUCT IMPROVEMENT:

Customers with specific service plans (e.g. international plan, voicemail) are more likely to churn. Evaluate the possibility of introducing new features or bundling options to retain these customers.

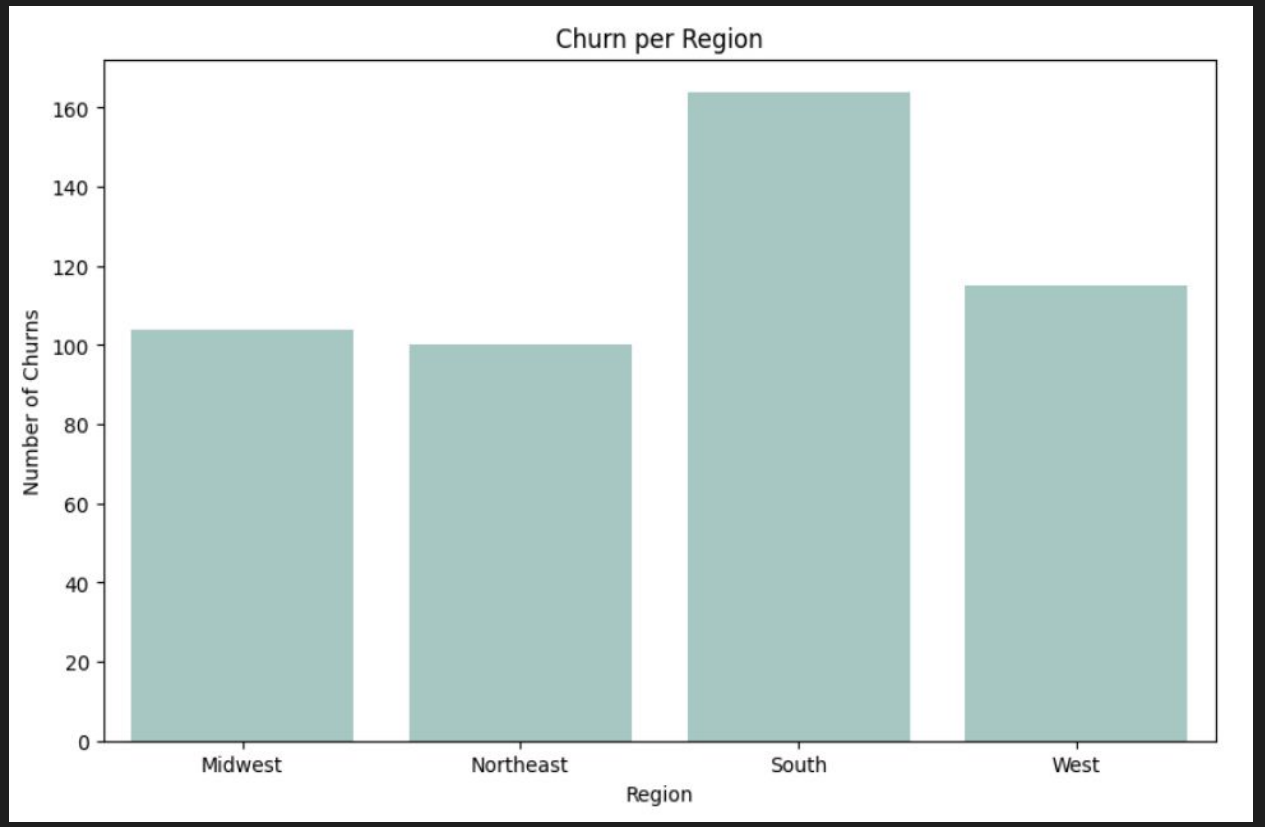
OPTIMIZE PRICING MODELS:

Offer flexible pricing models or discounts for long-term customers or those with high usage (e.g., total minutes or charges)

KEY VISUALS

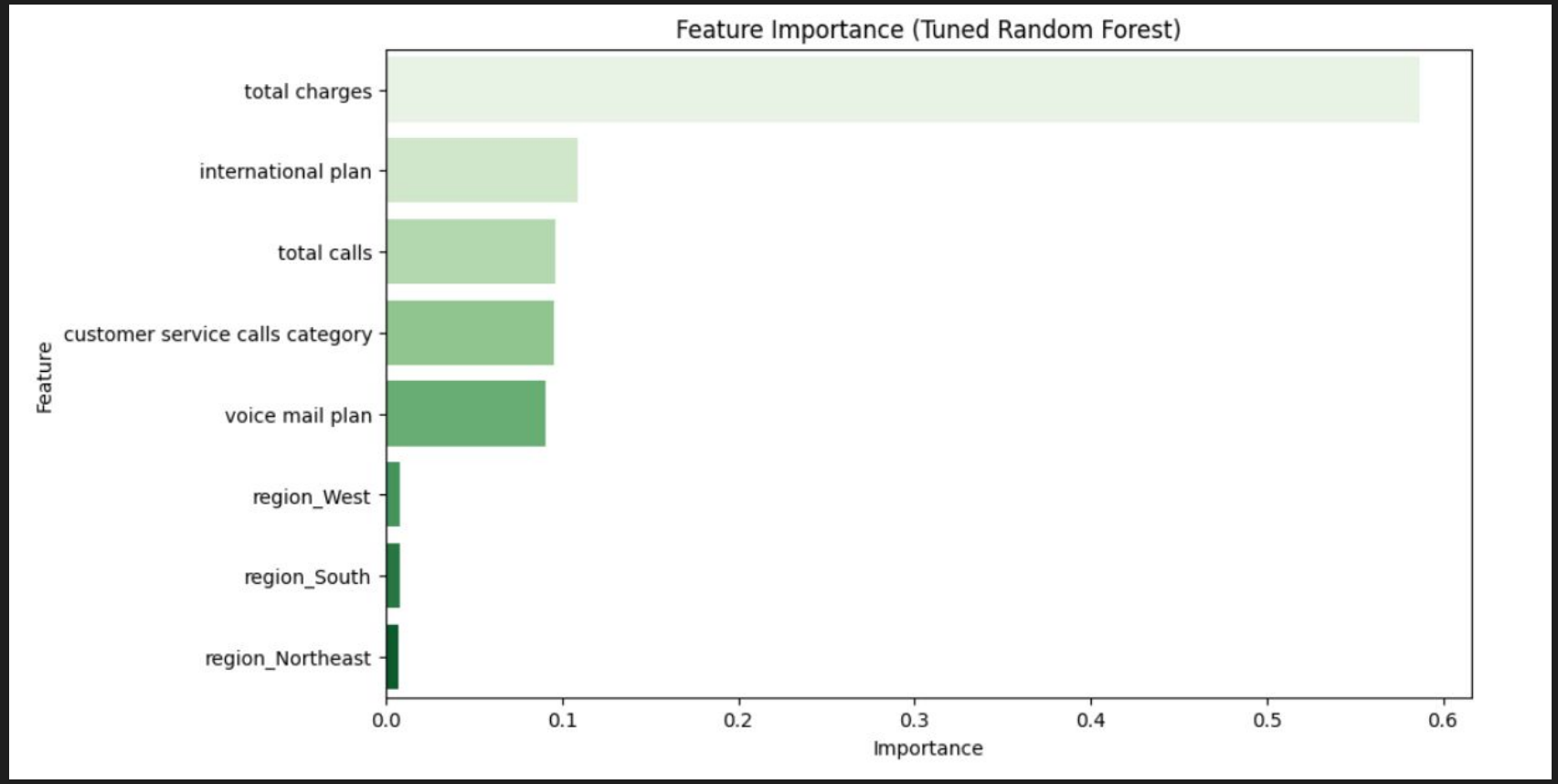
```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 3333 entries, 0 to 3332
Data columns (total 21 columns):
 #   Column                  Non-Null Count  Dtype
---  ------                  --------------  -----
 0   state                   3333 non-null   object
 1   account length          3333 non-null   int64
 2   area code               3333 non-null   int64
 3   phone number            3333 non-null   object
 4   international plan      3333 non-null   object
 5   voice mail plan         3333 non-null   object
 6   number vmail messages   3333 non-null   int64
 7   total day minutes       3333 non-null   float64
 8   total day calls         3333 non-null   int64
 9   total day charge        3333 non-null   float64
 10  total eve minutes       3333 non-null   float64
 11  total eve calls         3333 non-null   int64
 12  total eve charge        3333 non-null   float64
 13  total night minutes     3333 non-null   float64
 14  total night calls       3333 non-null   int64
 15  total night charge      3333 non-null   float64
 16  total intl minutes      3333 non-null   float64
 17  total intl calls        3333 non-null   int64
 18  total intl charge       3333 non-null   float64
 19  customer service calls  3333 non-null   int64
 20  churn                   3333 non-null   bool
dtypes: bool(1), float64(8), int64(8), object(4)
memory usage: 524.2+ KB
```

This is the dataset used for analysis and modeling



This shows churn per region. It could be explored for customized solutions for high churn regions

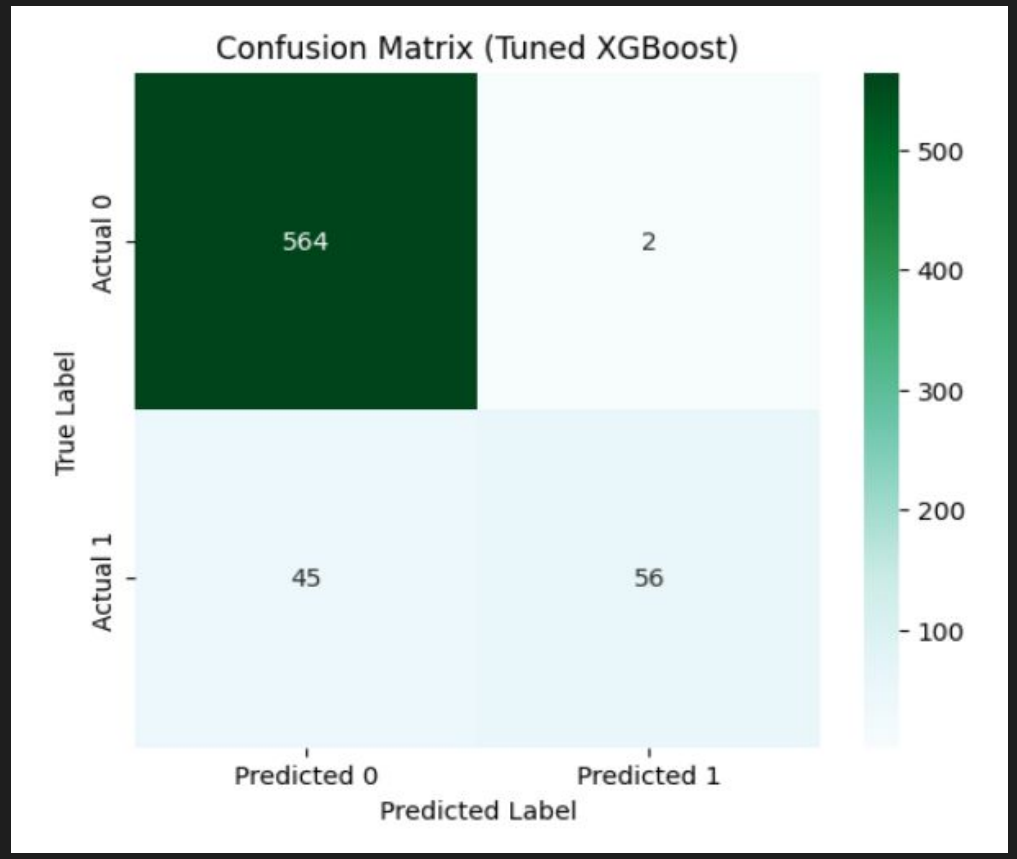
KEY VISUALS



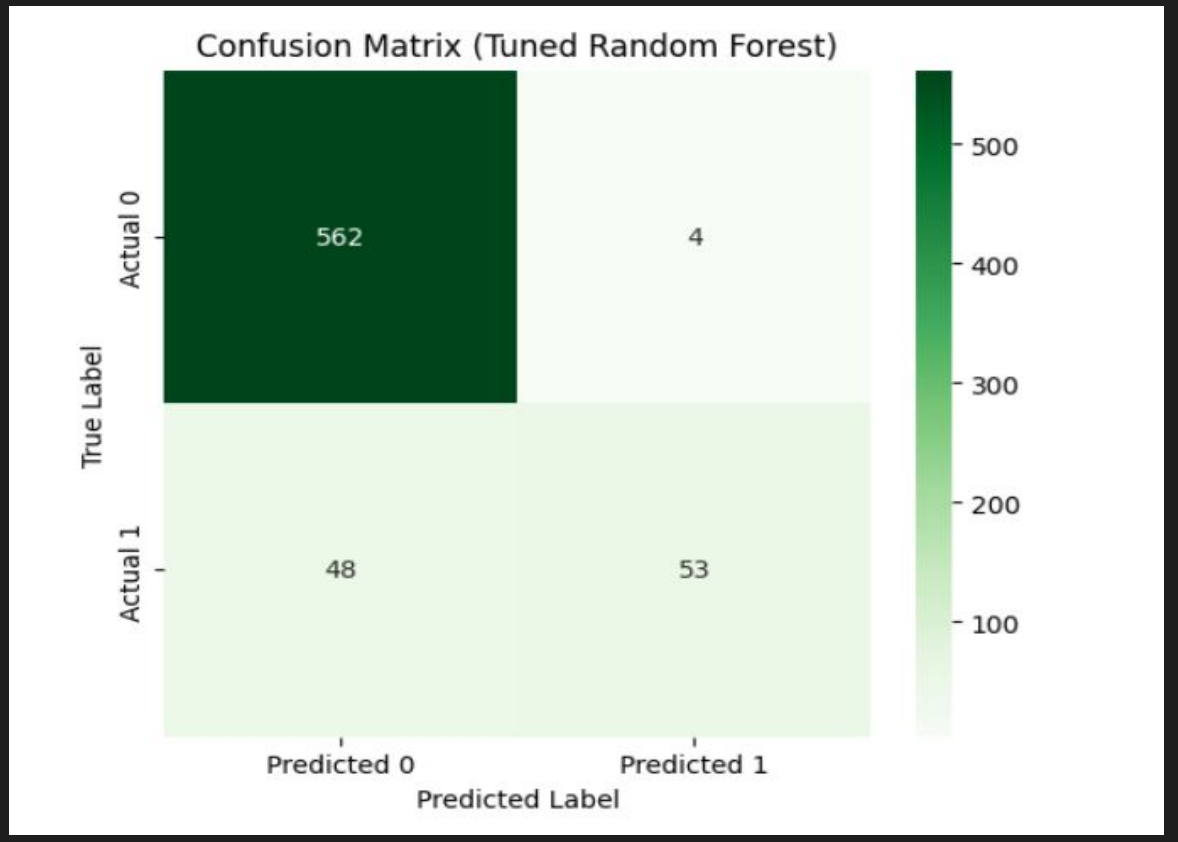
Bar Plot showing all features and how they affect churn

KEY VISUALS

Best models:



Confusion Matrix for Tuned XGBoost Model



Confusion Matrix for Tuned Random Forest Model

NEXT STEPS:

MODEL DEPLOYMENT:

Deploy the tuned models to identify high-risk customers on an ongoing basis.

CUSTOMER SEGMENTATION:

Create a customer segmentation strategy based on churn risk and develop personalized retention strategies..

MONITOR AND ADJUST:

Monitor churn rate regularly and update models based on new data.

KEY ANALYSIS CHALLENGES

IMBALANCED DATASET

There was data imbalance between churn and non-churn customers.

Churn was 14% and non churn was 86%

PRECISION RECALL TRADEOFF

High Recall, Low Precision: The model catches most churners but may waste resources on false alarms.

High Precision, Low Recall: The model is more accurate but may miss potential churners, limiting intervention opportunities.

RESOURCE CONSTRAINTS FOR DEPLOYMENT

Deploying and integrating the model requires resources and effort.

Without proper resources, the model may not be fully implemented, limiting its effectiveness.

THANK YOU