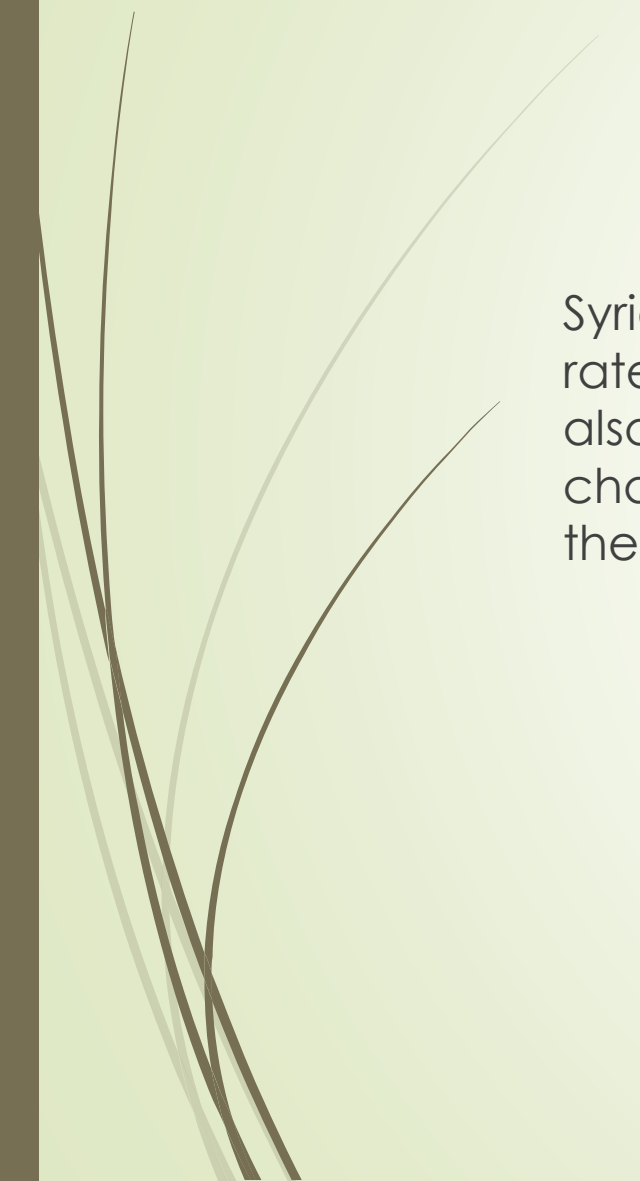# PHASE3 PROJECT : Customer Churn Prediction - SyriaTel

Student : Chepchirchir Hanifa

Technical Mentor : Nikita Njoroge

OVERVIEW

SyriaTel, a leading telecom provider, is experiencing a high customer churn rate—losing nearly 1 in 7 customers annually. This not only impacts revenue but also increases the cost of acquiring new users. To address this business challenge, we analyzed customer behavior data to predict churn and identify the key drivers behind it.

Problem statement

- SyriaTel is losing nearly 15% of its customers annually, impacting revenue and customer lifetime value. The company currently lacks a proactive system to identify customers at risk of churning.

- This project uses customer usage and service data to build a predictive model, helping the marketing and retention teams intervene early with personalized actions to reduce churn and improve loyalty.

- **Objective:**

- SyriaTel is facing customer churn, which directly impacts revenue. The Customer Retention and Marketing Team wants to understand and predict which customers are most likely to churn.

- Build a classification model that predicts whether a customer will churn using features like service usage, international plan status, voicemail activity, and customer support interactions.

- Key Questions to Answer:

- Can we predict churn based on customer behavior?

- What are the most influential features in predicting churn?

- How can SyriaTel use these insights to reduce churn?
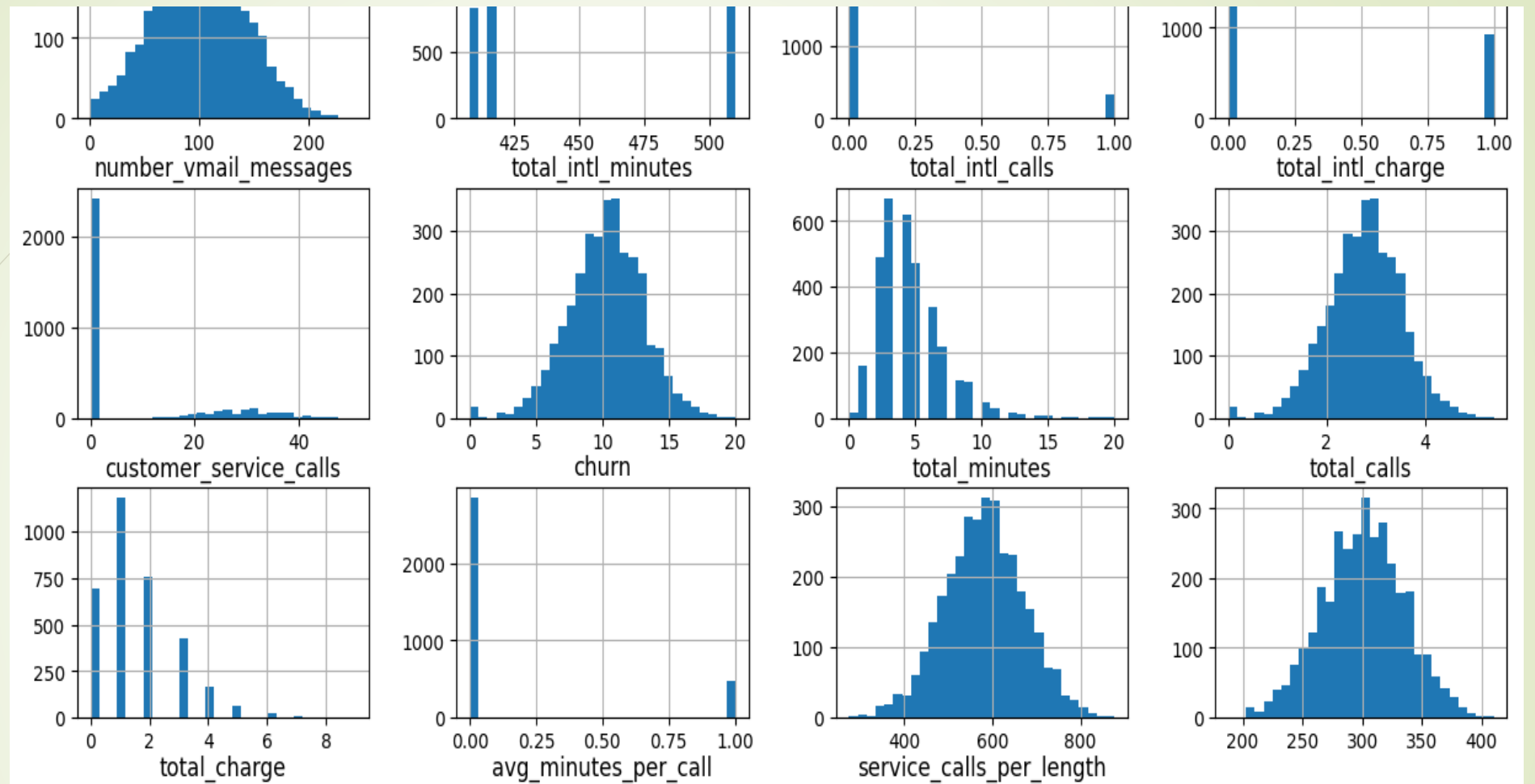
Plotting distribution of numerical features

Identifying which features that could work well with the target which is churn from the visuals features like customer_service_calls will work well because it shows clear distinctions
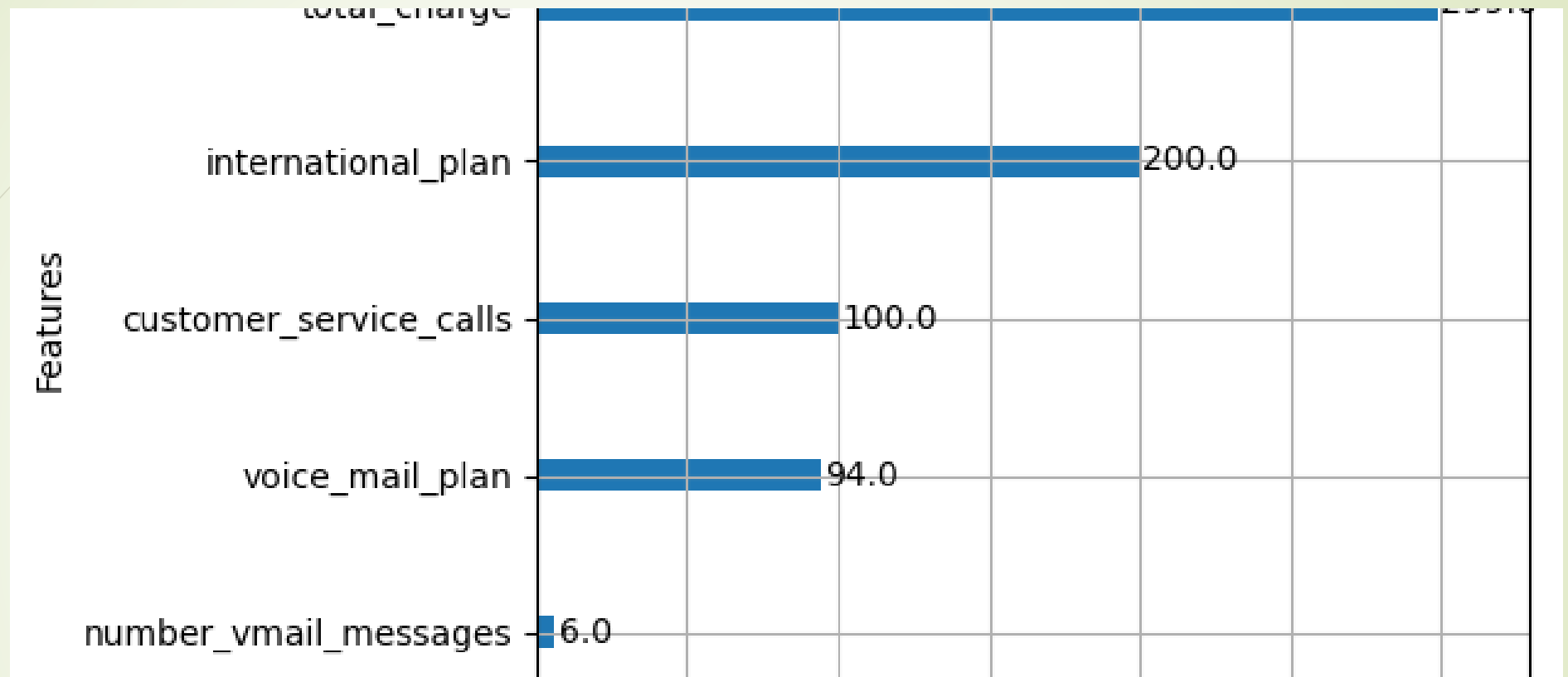
Top predictive features for churn

These features performs well with our model since it is balanced in predicting customer churn
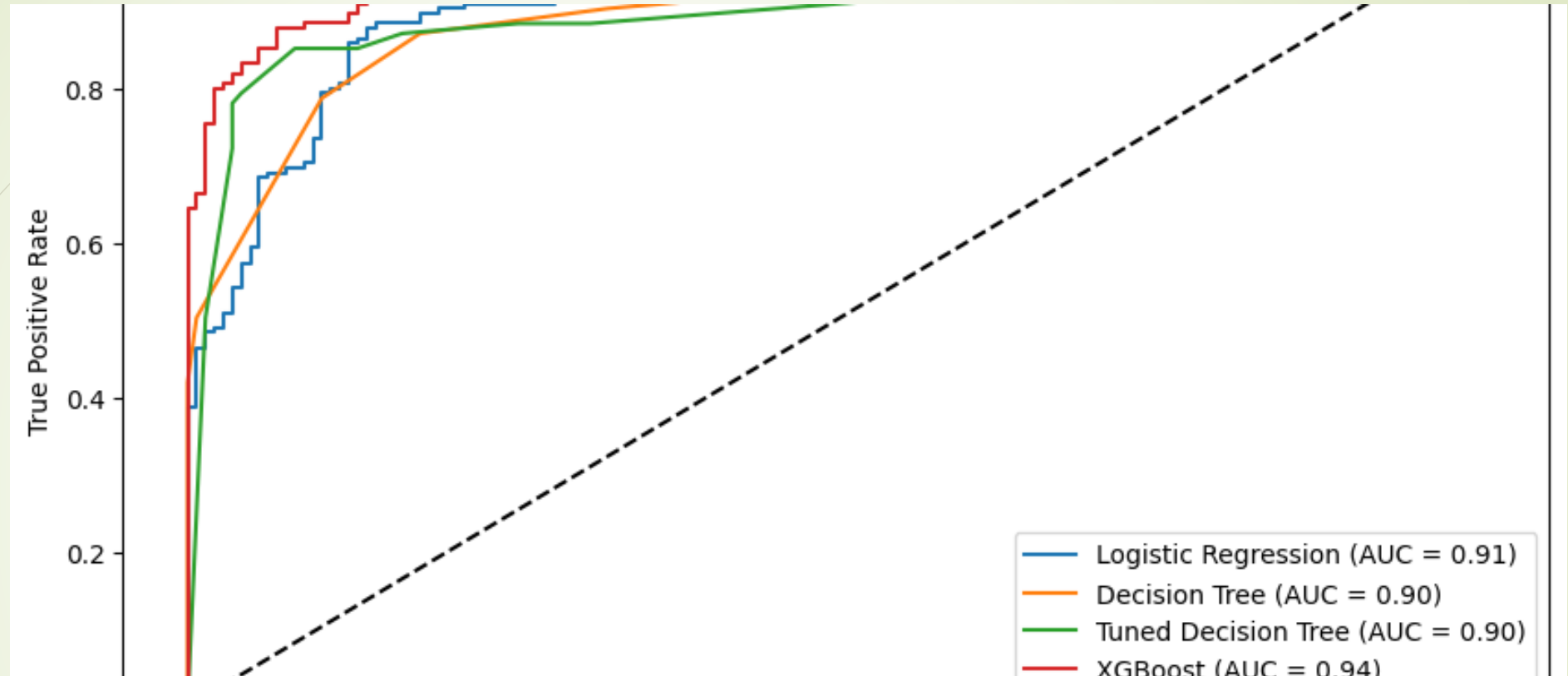
ROC Curves for Different Models

The Roc curves suggests that XGBoost (AUC=0.94) as the top-performing model for churn prediction compared to the other curves .These results indicate XGBoost would most effectively identify at-risk customers while minimizing false alarms.

Project objectives

- Can we predict churn based on customer behavior?

- Yes Behavioral data such as **call patterns, service usage, and interactions with customer support** contains strong signals about whether a customer is likely to leave.

- What are the most influential features in predicting churn?

- Customer Service Calls, International Plan, Total Day Minutes, Total International Calls/Minutes/Charges etc

- Churn is highly influenced by **service experience**, **plan types**, and **usage behavior**. These features help telecom companies **identify at-risk customers** and **tailor retention strategies** more effectively.

- How can SyriaTel use these insights to reduce churn?

- SyriaTel can **translate model insights into targeted business actions** to proactively retain customers. Here's how:

- High customer service call frequency strongly correlates with churn.

- Customers on international plans are more likely to churn.

- High day-time usage may signal price-sensitive users.

- Predictive models can accurately classify churn-prone customers.

- By using these insights, SyriaTel can move from **reactive customer service** to **proactive churn prevention**, increasing customer satisfaction and long-term revenue retention.

Key findings

- **Frequent Customer Service Calls**
  ➤ Improve agent training, resolve issues faster, and follow up with dissatisfied users.

- **International Plan Users**
  ➤ Reevaluate pricing, offer tailored bundles, and enhance call quality.

- **No Voice Mail Plan**
  ➤ Promote voicemail through upsells, free trials, and bundled plans.

- **High Daytime Charges**
  ➤ Recommend personalized plans, send usage alerts, and offer loyalty discounts.

- **Cost Sensitivity**
  ➤ Ensure transparent billing and proactively support price-sensitive customers.

.

Conclusions

- This project successfully explored customer churn prediction using the SyriaTel dataset. After extensive preprocessing, feature engineering, model experimentation, and hyperparameter tuning, the XGBoost Classifier emerged as the best-performing model with:

- Accuracy: 94%

- F1-Score: 0.77

- ROC AUC: 0.84

- XGBoost demonstrated a solid balance between precision and recall, making it ideal for detecting potential churners while minimizing false alarms. Decision Tree and Random Forest also showed competitive performance but with lower precision and slightly reduced generalizability.

- Key features influencing churn included:

- Customer Service Calls

- Total Call Minutes

- Voicemail Plan

- International Plan

- By identifying churn-driving factors, optimizing classification models, and implementing customer retention strategies, this project provides valuable insights for SyriaTel. Next steps focus on refining prediction accuracy, enhancing business interventions, and improving customer engagement.

Recommendations

- Targeted Retention Campaigns
- Focus marketing and retention efforts on customers flagged by the model, especially those who frequently call customer service and lack a voicemail plan.
-

  Service Quality Monitoring
- Monitor high call volumes to customer support as early churn indicators. Enhancing support responsiveness could help retain such customers.
-

  Product Bundling
- Encourage customers to subscribe to voicemail or international plans, as they are associated with lower churn rates.

Next steps

- **Feature Engineering Enhancements**: Add tenure, call pattern trends, and customer demographics

- **Hyperparameter Tuning**: Use Bayesian optimization for improved performance

- **Threshold Optimization**: Adjust probability thresholds for better balance between precision and recall

- **Deployment**: Develop an API or integrate into CRM systems for real-time predictions