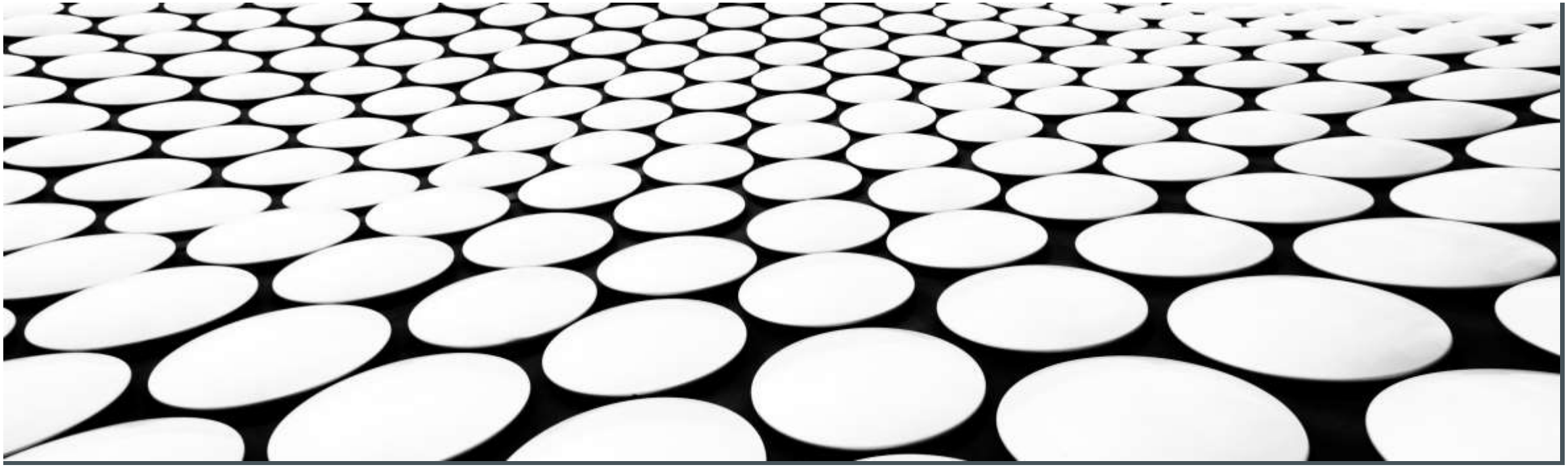
SYRIATEL CUSTOMER CHURN PREDICTION

LEILA NYAMBURA, DSF-PT07P0 - PHASE 3 PROJECT (AUGUST 2024)



OUTLINE

- Overview
- Business and Data Understanding
- Modeling
- Evaluation
- Identified Key Churn Drivers
- Recommendations
- Next Steps

OVERVIEW

Project Objective: To predict customer churn at SyriaTel and identify factors leading to customer churn, enabling targeted interventions to reduce churn and improve customer retention.

BUSINESS AND DATA UNDERSTANDING

Business Problem: SyriaTel has been experiencing customer churn. To address this, the project aims to predict which customers are likely to leave and understand the factors contributing to churn.

Key Stakeholders:

- **Customer Relationship Management (CRM) Team:** to use the churn prediction insights to design retention strategies.
- **Marketing Team:** to design and execute personalized campaigns based on the predicted churn risk and identified churn factors.
- **Customer Service/Support Team** to proactively reach out to at-risk customers, improving service quality and addressing issues that may lead to churn.

Data Overview: The dataset includes customer information, usage patterns, service plans, and customer service interactions.

MODELING

Approach

I used classification modeling to predict customer churn. Classification models are useful because they help categorize customers into 'likely to churn' or 'not likely to churn’.

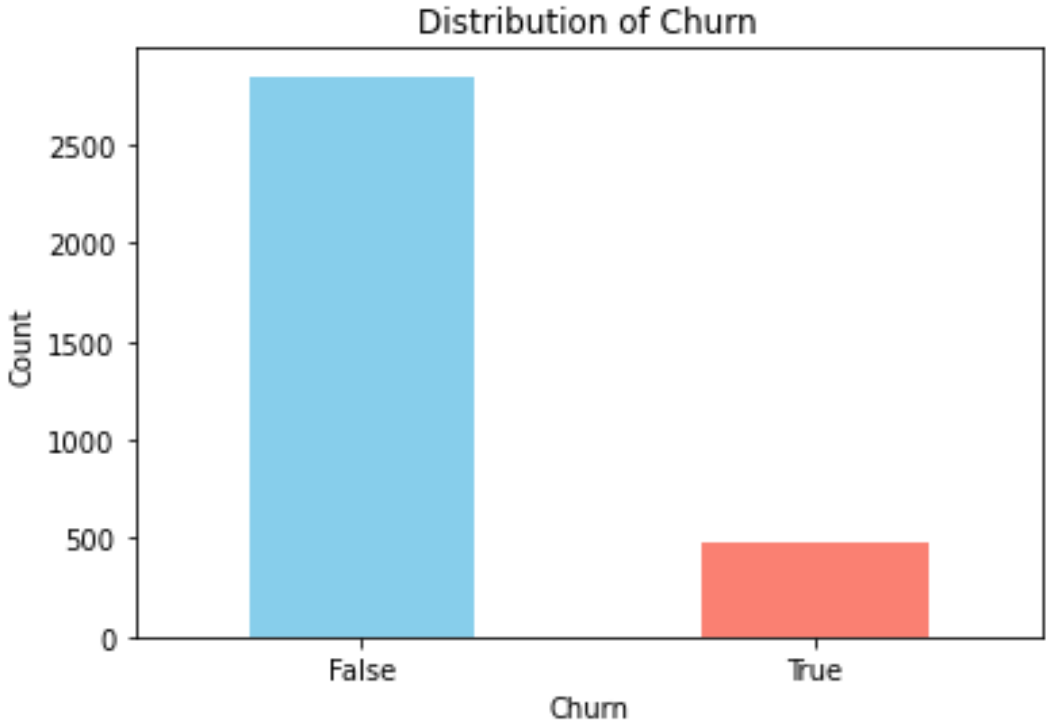
Models Used:

- **Baseline Model:** Logistic Regression (Simple and interpretable).
- **Decision Tree Model:** Captures more complex patterns but may overfit.
- **Tuned Decision Tree:** Optimized to reduce overfitting and improve generalization.

CHURN DISTRIBUTION

The image to the right shows the churn distribution from the dataset used.

- False (blue): indicates customers will not leave SyriaTel.
- True (orange): indicates customers will leave SyriaTel.



EVALUATION

1. During the model evaluation the **key metrics** were:

- **Accuracy:** Percentage of correct predictions.
- **Precision:** Of the predicted churners, how many actually churned.
- **Recall:** Ability to identify actual churners (important for minimizing customer loss).
- **F1-Score:** Balance between precision and recall.

All these metrics are important in identifying the best performing model at predicting the customer churn. From the metrics, I looked at **model comparison** and below were my findings:

- **Logistic Regression:** Good accuracy but low recall, meaning it missed many churners.
- **Decision Tree:** High recall but overfitted, performing perfectly on training data.
- **Tuned Decision Tree:** Balanced model with reduced overfitting, making it the best choice for predicting churn.

Identified Key Churn Drivers

1.Total Charges (e.g., Total Day, Evening, Night, and International Charges Combined)

- Impact: High charges can indicate that customers are being charged more than they expect or can afford, leading to dissatisfaction and a higher likelihood of churn.

2.Customer Service Calls

- Impact: A high number of calls to customer service often signals unresolved issues or dissatisfaction, which can drive customers to leave.

3.International Plan Usage

- Impact: Customers who use international services may be more price-sensitive or have specific needs that, if unmet, can lead to churn. How often they use the plan and the associated charges can influence their satisfaction.

4.Voicemail Plan Usage

- Impact: Similar to international plan usage, customers with specific needs (e.g., voicemail) might churn if these needs aren't met satisfactorily or if they feel the cost isn't justified.

Identified Key Churn Drivers cont’

5.Total Minutes Used

- Impact: The total minutes used, whether during the day, evening, or night, can reflect how engaged a customer is with the service. Very high or very low usage might indicate potential churn, depending on whether the customer feels they are getting value for money.

6.Average Call Duration

- Impact: The length of calls might indicate the level of service interaction or satisfaction. Short calls might suggest poor engagement, while unusually long calls could indicate issues that aren't being resolved efficiently.

These features are typically identified as important in predicting churn because they directly relate to customer satisfaction, cost sensitivity, and service usage patterns. By addressing these areas, SyriaTel can target the factors most likely to lead to customer dissatisfaction and churn.

RECOMMENDATIONS

From my model findings these would be my recommendations to the SyriaTel stakeholders:

1. Implement Targeted Retention Strategies:

- Use the Tuned Decision Tree model to identify high-risk customers and design personalized retention offers.

2. Address Key Churn Drivers:

- Focus on reducing issues related to high service charges and frequent customer service calls.

3. Monitor and Refine the Model:

- Continuously update the model with new data and monitor its effectiveness to adapt to changing customer behaviors.

NEXT STEPS

As a starting point SyriaTel should focus on the next steps mentioned below:

1.For Targeted Retention Strategies:

- The CRM Team should develop personalized retention strategies (e.g., loyalty programs, discounts) for at-risk customers.
- The Customer Service/Support Team should proactively reach out to identified at-risk customers to offer improved service and address any issues.

2.For Leveraging Model Insights:

- The Marketing Team should use the identified churn drivers to craft personalized marketing campaigns addressing specific customer pain points.
- The CRM Team should collaborate with other departments to address systemic issues driving churn, such as network reliability or pricing concerns.

3.For Continuous Monitoring and Refinement:

- Regularly retrain the churn prediction model using the latest data to maintain its accuracy.
- Track the success of retention efforts through metrics like customer retention rates and satisfaction scores.
- Consider integrating additional data sources (e.g., customer feedback, social media sentiment) to enhance the model's predictive power.

THANK YOU!

For any questions and further analysis, kindly reach out to me on:
Email: leilanyambura@students.moringaschool.com/
Github: Leila-Nyambura