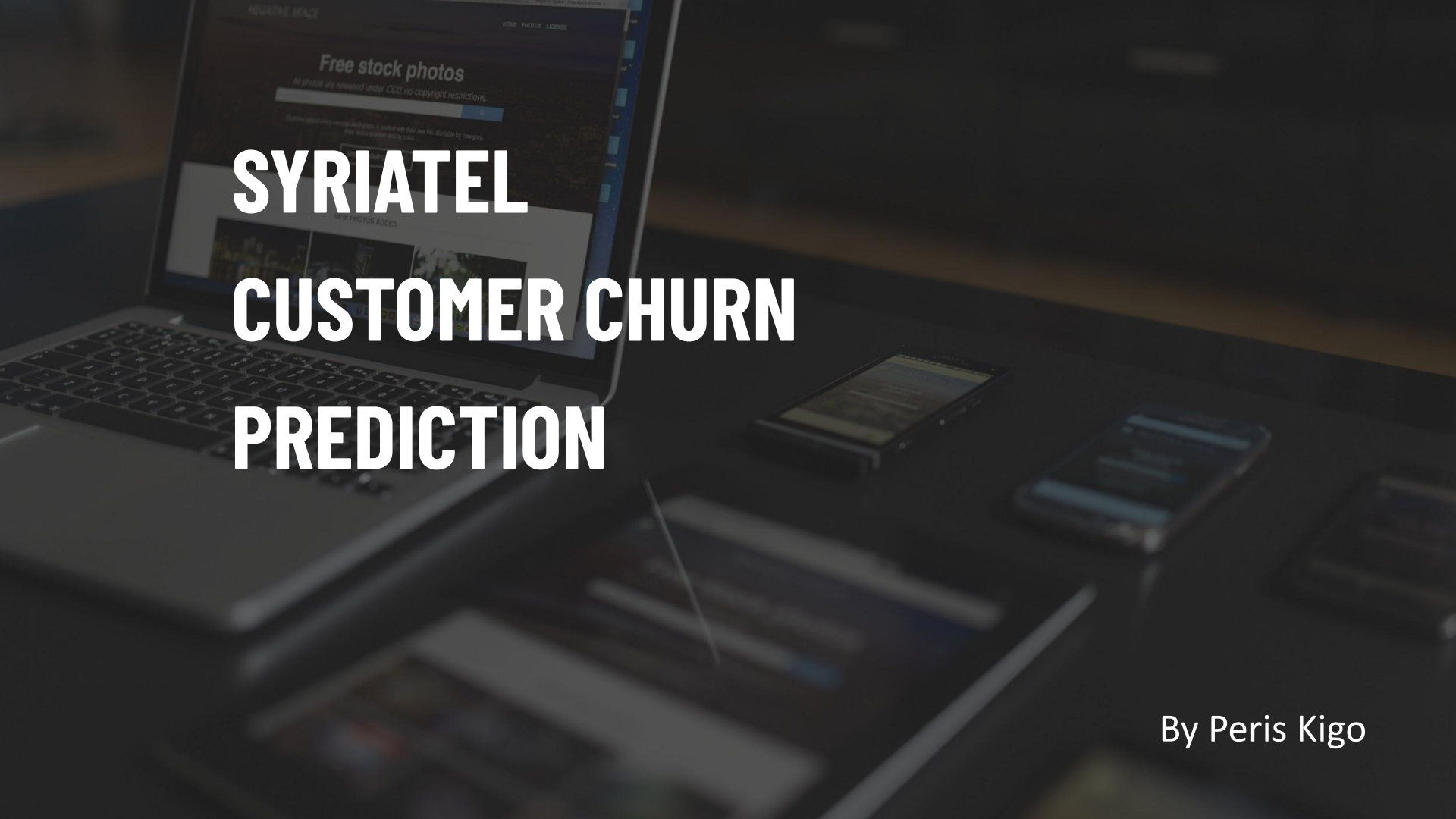


SYRIATEL CUSTOMER CHURN PREDICTION

By Peris Kigo

Overview

In this project, we analyzed SyriaTel's customer data to explore what factors contribute to customer churn. Churn is when customers stop using the company's services, and it poses a major financial risk for telecom businesses.

Objectives.

1. To identify the factors that significantly influence churn.
2. To build and validate a predictive model that classifies whether a customer is likely to churn or likely to stay.
3. Provide insights and recommendations to reduce churn and improve retention.

Business Understanding

In the highly competitive telecommunications industry, retaining customers is just as important as acquiring new ones.

- Problem: Syriatel is experiencing customer churn that affects revenue and long-term growth.
- Impact: Losing customers is costly, since retaining a customer is much cheaper than acquiring a new one

This analysis is designed to inform Syriatel's customer management strategy, minimize losses, and enhance loyalty.

Data Understanding

The dataset contains records for 3,333 customers, covering both usage patterns and service details.

Features include:

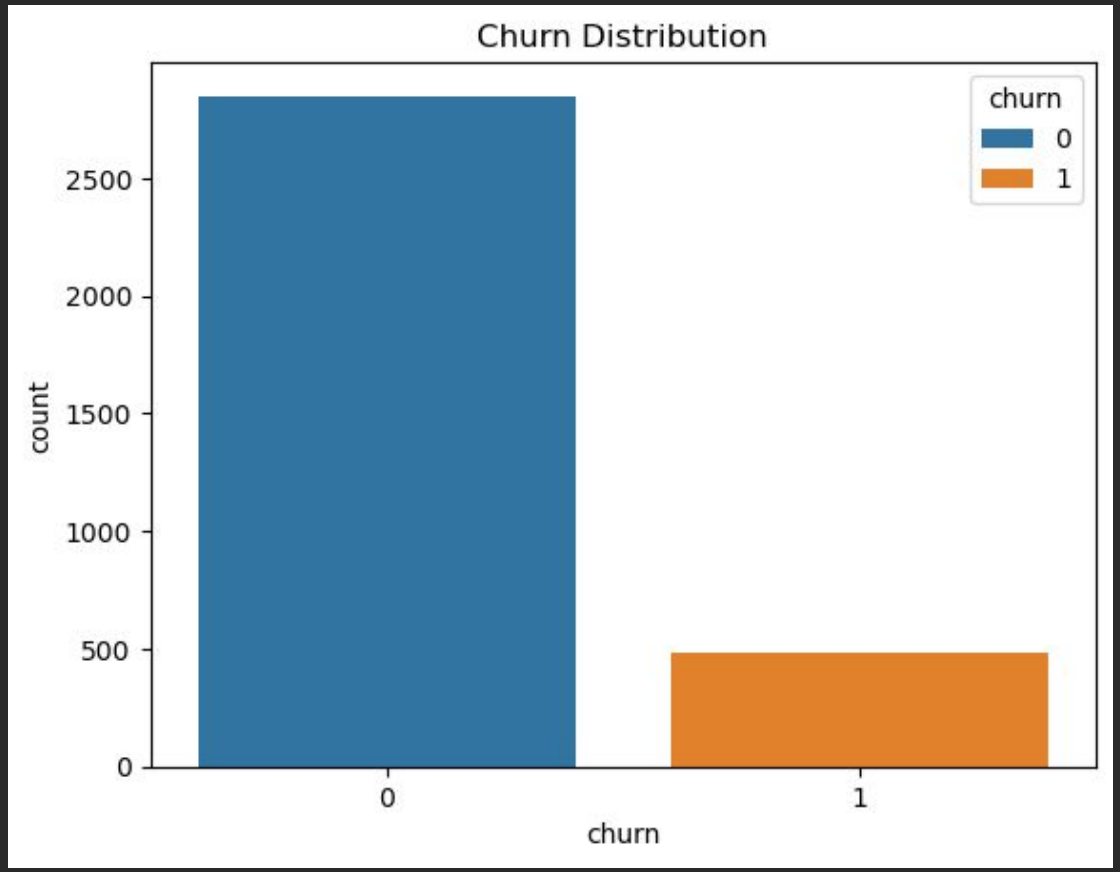
1. Demographics: State, area code, phone number.
2. Usage data: Day, evening, night, and international call minutes, calls, and charges.
3. Plans: Whether a customer has an international or voice mail plan.
4. Customer service interactions: Number of service calls.
5. Target: Churn (1 = churned, 0 = retained).

Data Analysis

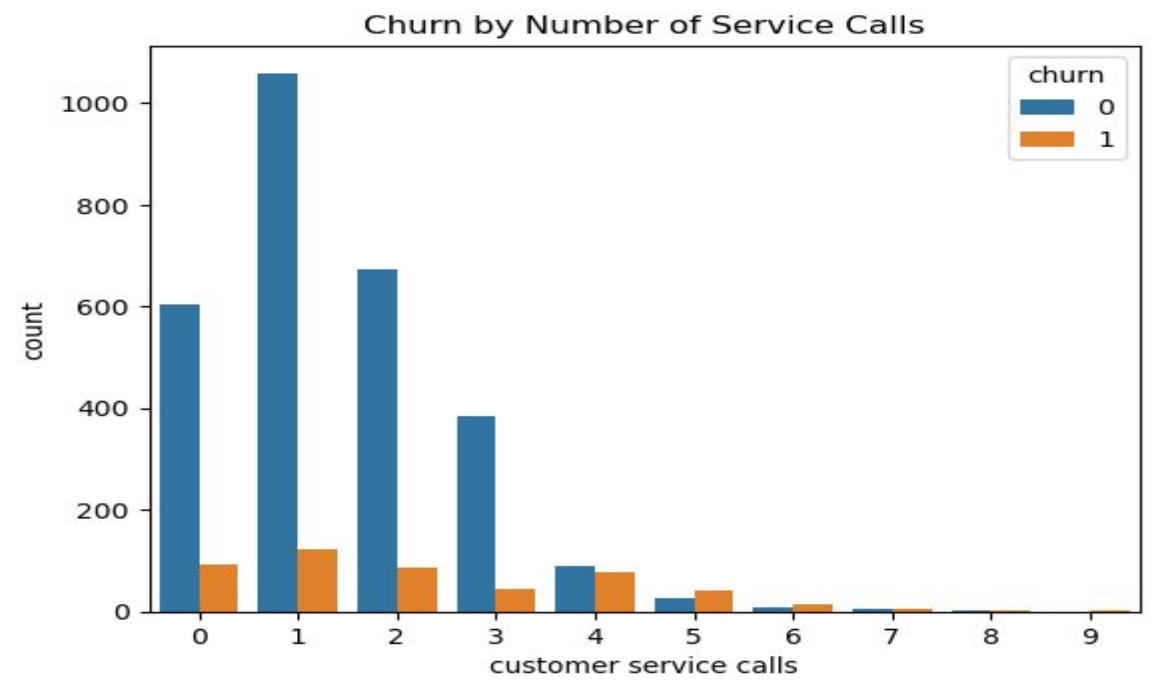
Before analysis we first did some data cleaning where we:

- Checked for missing values (none found).
- Removed irrelevant identifiers (e.g. phone number).
- Converted categorical variables into numerical (dummy encoding).
- Standardized column naming for consistency.
- Verified class balance and applied class_weight="balanced" in modeling.

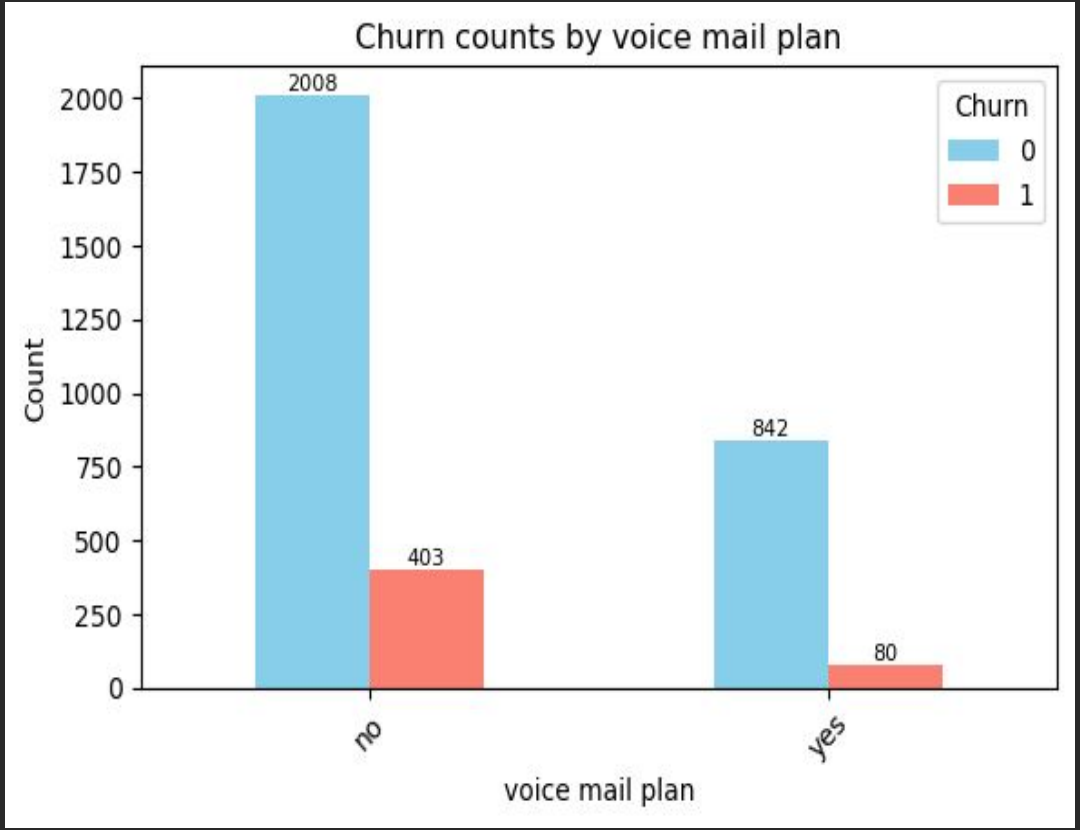
After cleaning we proceeded to Exploratory Data Analysis(EDA)



Out of 3,333 customers, 2,850 (85.5%) remained, while only 483 (14.5%) decided to leave.

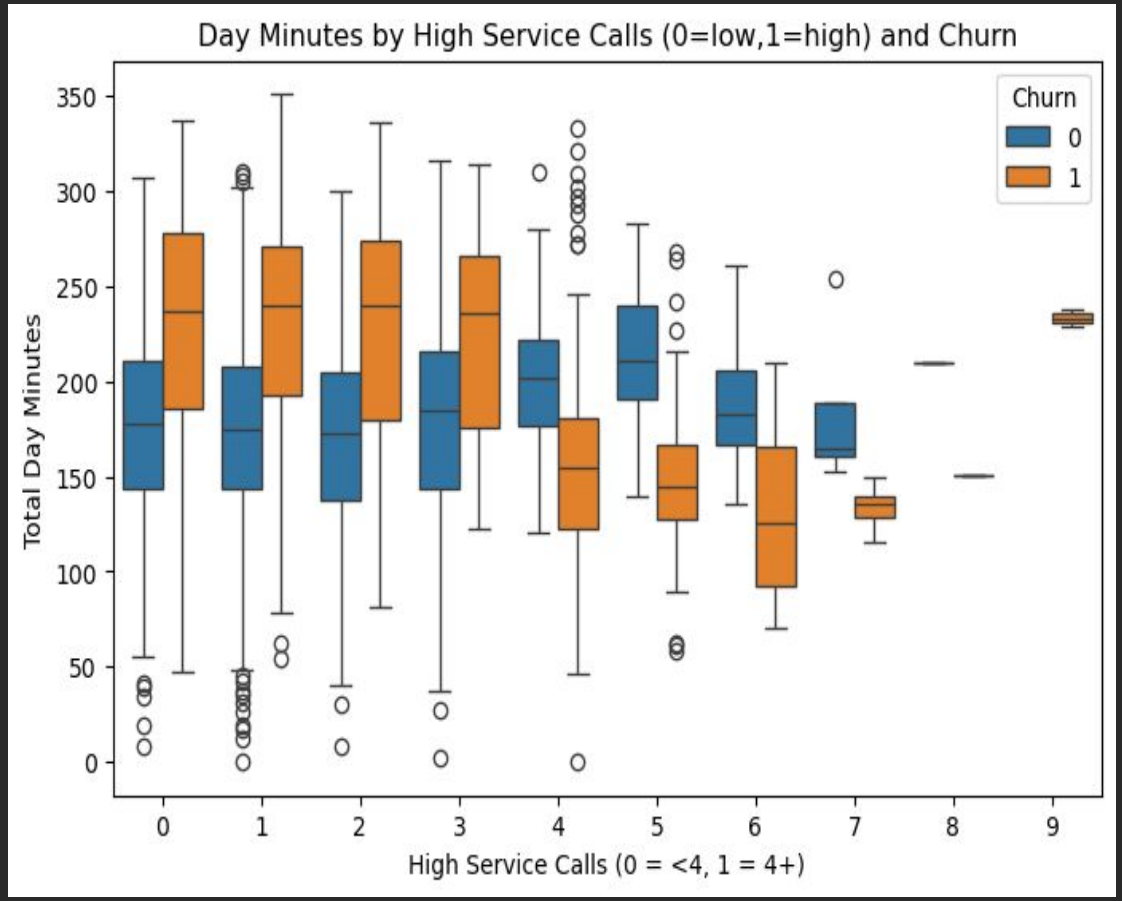


Most customers make less than 3 calls,however, customers with more than 1 service call have a noticeable spike in churn.



Customers with a Voice Mail Plan tend to churn less compared to those without it.

This suggests that the voicemail service may add perceived value and satisfaction, helping retain customers



Customers flagged with high_service_calls (4+) show wider spread in day minutes, and a larger share of churn within that group.

This indicates that both high usage and repeated service contacts together are correlated with churn.

Modeling Approach

We built multiple machine learning models to predict churn:

1. **Baseline Model:** Logistic Regression to establish a benchmark.
2. **Decision Tree:** To understand feature-based splits and churn rules.
3. **Random Forest:** An ensemble method to improve predictive accuracy.
4. **SMOTE (Synthetic Minority Oversampling):** Used to address class imbalance and improve recall.
5. **Hyperparameter Tuning:** Optimized Random Forest to achieve best results

Evaluation

The models were evaluated using accuracy, precision, recall, and F1-score.

- **Logistic Regression (baseline):** Good accuracy but struggled with recall due to class imbalance.
- **Decision Tree:** Improved interpretability but prone to overfitting.
- **Random Forest:** Best performing model with 94% accuracy after tuning, and strong balance between precision and recall.

Key predictors of churn included:

- International plan.
- Number of customer service calls.
- Total day minutes.

Recommendations.

Based on our findings, we recommend the following strategies for Syriatel:

1. **Improve Customer Service:** Focus on faster resolution of issues to reduce dissatisfaction from repeat calls.

2. **Review International Plans:** Offer better pricing or incentives for international users who have higher churn rates.

3. **Target churning Customers:** Use the churn prediction model to flag high-risk customers for retention campaigns.

4. **Develop Daytime-Friendly Packages:** Cater to heavy daytime users who are more likely to churn.

Next steps

1. Deploy the churn prediction model into Syriatel’s customer relationship management system.

2. Continuously retrain the model on new data to keep predictions accurate.

3. Design targeted retention campaigns based on model outputs

THANK YOU

Do you have
any questions?