

Project Proposal: Customer Churn Prediction for SyriaTel

Business Understanding

Business Overview:

SyriaTel, a leading telecommunications company, is determined to tackle revenue loss attributed to customer churn. The primary goal is to construct a predictive model that can discern whether a customer is likely to terminate their services, framing the task as a binary classification challenge. The key stakeholders driving this initiative are SyriaTel's management and key stakeholders, singularly focused on curbing customer churn and augmenting revenue retention. These objectives are paramount in sustaining and advancing SyriaTel's standing within the fiercely competitive telecommunications industry.

Business Objectives:

1. Develop a classifier to predict customer churn.
2. Identify patterns or features indicative of potential churners.
3. Enable SyriaTel to implement targeted retention strategies.

Data Understanding

Data Inventory:

The dataset encompasses historical customer interactions, behaviors, and churn events for SyriaTel's telecommunications services.

Assumptions:

- Data provided is assumed to be accurate and up-to-date.
- Data anomalies are expected to be minimal.
- The data is unlabeled for unsupervised learning.

Data Mining Goals:

- Identify customers with higher churn propensity.
- Segment customers based on usage patterns.
- Identify key factors influencing churn.

Data Preparation

- Importing necessary libraries and loading the data.
- Cleaning and preprocessing the data, handling missing values, duplicates, etc.
- Engineering relevant features crucial for churn prediction.
- Converting relevant variables into appropriate formats.

Exploratory Data Analysis (EDA)

Univariate Analysis:

1. Uncovering Patterns in Customer Behavior:

This involves examining various aspects of customer behavior to identify trends and patterns that may be indicative of churn. Key steps may include:

- Call Usage Patterns: Analyzing the distribution of call durations, frequencies, and types (e.g., voice, text, data) to understand how customers engage with SyriaTel's services. This can provide insights into high or low activity periods.
- Usage Intensity: Exploring the intensity of service usage (e.g., high data usage, long call durations) to discern if there are groups of customers with distinct usage patterns.
- Service Preferences: Investigating which services (e.g., international calls, specific features) are more popular among customers and whether there are correlations with churn behavior.

2. Analyzing Usage Patterns by Time of Day, Week, etc.:

Understanding how usage patterns vary over time can provide valuable insights into customer behavior. This involves:

- Temporal Trends: Examining how usage changes during different times of the day, days of the week, or even across seasons. For example, are there more calls during evenings or weekends?
- Seasonal Variation: Identifying any seasonal trends or cyclical patterns in usage that may correlate with churn behavior.
- Peak Usage Times: Determining when customers are most active and if there are specific time frames associated with increased likelihood of churn.

3. Investigating Trends in Customer Demographics:

Analyzing customer demographics can help identify if certain groups are more prone to churn. This may include:

- Age and Gender Analysis: Assessing whether age or gender influence churn rates. For instance, do younger customers exhibit different churn behavior compared to older ones?
- Geographical Analysis: Investigating whether customers from specific regions or states are more likely to churn. This may provide insights into regional preferences or competitive influences.

- Plan Preferences: Understanding if customers with specific plan types (e.g., international plans, data-heavy plans) have different churn rates.

4. Additional EDA Observations:

Depending on the available data, consider exploring other relevant factors that may influence churn behavior. This could include:

- Customer Complaints: If available, examining customer feedback or complaint data to identify recurring issues that may lead to churn.
- Customer Tenure: Analyzing how the length of time a customer has been with SyriaTel correlates with their likelihood of churning.
- Customer Feedback or Surveys: If survey data is accessible, leveraging sentiment analysis to gauge customer satisfaction and its relation to churn.

By conducting a thorough EDA, we aim to gain comprehensive insights into customer behavior, enabling us to identify key factors that influence churn propensity. This knowledge will form the foundation for constructing an accurate and effective customer churn prediction model.

Modeling

- Employing machine learning algorithms (e.g., Logistic Regression, Random Forest) for churn prediction.
- Evaluating model performance using appropriate metrics (accuracy, precision, recall, F1-score).
- Implementing cross-validation to ensure model robustness.

Observations and Conclusions

- Providing insights into customer segments with higher churn propensity.
- Identifying key features influencing churn decisions.
- Evaluating model performance and its implications.

Recommendations

- Implement targeted retention strategies for identified high-risk customer segments.
- Enhance customer service during high-churn periods.
- Explore personalized offerings to incentivize customer retention.

Expected Impact

- Reduced customer churn leading to increased revenue retention.
- Enhanced customer satisfaction and loyalty.
- Strengthened competitive position in the telecom industry.

