Project Abstract

Customer Churn Prediction

TEAM MEMBERS

Senthil Kumar J - 2021115098

Shruthi B - 2021115106

Yalini Kumar - 2021115122

Yuthikshaa M - 2021115123

Kattu Bava K S - 2021115308

PROBLEM DEFINITION

In the highly competitive business landscape, customer retention is paramount for sustainable growth and profitability. The issue of customer churn, where valuable clients discontinue their relationship with a company, looms as a critical concern. To safeguard revenue streams and foster long-lasting customer relationships, the development of an effective customer churn protection strategy is imperative.

Customer churn is influenced by a multitude of factors, including product satisfaction, pricing strategies, customer service quality, and external market dynamics. It not only results in immediate revenue loss but also necessitates increased efforts and resources to acquire new customers. Hence, the core objective of this project is to create a comprehensive customer churn protection system. This system will harness data-driven insights, proactive measures, and cutting-edge technologies to predict potential churners accurately, implement targeted retention strategies, and continuously refine approaches based on real-time feedback. By tackling the challenge of customer churn head-on, this initiative aims to enhance customer retention rates and secure long-term business success.

DATASET CONTENT

1. **Customer Information:**
    - Customer ID or unique identifier
    - Customer name (if applicable)
    - Customer contact information (email, phone number, address)
    - Customer demographics (age, gender, location)

2. **Usage and Behavior Data:**
    - Duration of customer relationship (tenure)
    - Frequency of product/service usage or purchases
    - Recency of interactions or purchases
    - Average transaction or purchase amount
    - Historical usage patterns (e.g., login frequency, click-through rates)

3. **Billing and Payment Information:**
    - Monthly subscription cost
    - Payment method (credit card, PayPal, etc.)
    - Billing cycle information
    - Payment history (missed payments, late payments)

4. **Customer Support and Interaction History:**
    - Number of customer service inquiries or tickets raised
    - Response times to customer inquiries
    - Customer feedback or satisfaction scores
    - Number of interactions with customer support

5. **Product/Service Features:**
    - Features or attributes related to the product or service being offered (e.g., plan type, features included)
    - Product usage data (e.g., app usage, feature adoption)
6. **Competitor Information (if applicable):**
    - Presence of competitors in the market
    - Competitor pricing and offerings

7. **Promotional and Marketing Data:**
    - Participation in marketing campaigns or promotions
    - Offers or discounts applied to the customer's account
    - Marketing channel source (e.g., online ads, referrals)

8. **Churn Label:**
    - A binary indicator (1 or 0) indicating whether the customer churned (1) or not (0) during a specified time period.

9. **Time-Related Data:**
    - Timestamps or dates of customer interactions and events (e.g., sign-up date, last login date)
    - Churn prediction period (the time frame for which churn is being predicted)

10. **Additional Features (if available and relevant):**
    - Customer sentiment or sentiment analysis from customer reviews or feedback
    - Social media activity related to the company or product
    - Customer referrals and referrals' behavior

DESIGN THINKING

1. **Data Collection and Preprocessing:**
    - Gather historical customer data, including features relevant to churn prediction (as mentioned in the previous answer).
    - Clean and preprocess the data, handling missing values, outliers, and data transformations as needed.

2. **Feature Engineering:**
    - Select and engineer relevant features from the dataset to capture customer behavior and characteristics effectively.
    - Create target labels indicating whether customers have churned during a specific time frame.

3. **Data Splitting:**
    - Split the dataset into three subsets: a training set, a validation set, and a test set. Common splits are 70% for training, 15% for validation, and 15% for testing.

4. **Model Selection:**
    - Choose an appropriate machine learning algorithm for customer churn prediction. Common algorithms include logistic regression, decision trees, random forests, gradient boosting, support vector machines, and neural networks.
    - Consider ensemble methods or stacking models for improved accuracy.

5. **Model Training:**
    - Train the selected model(s) using the training dataset.
    - Tune hyperparameters to optimize model performance using the validation dataset.

6. **Model Evaluation:**
    - Evaluate the trained model(s) using appropriate evaluation metrics, such as accuracy, precision, recall, F1-score, ROC-AUC, and confusion matrix.
    - Adjust the model's threshold (if applicable) to balance precision and recall based on business needs.

7. **Model Interpretability (Optional):**
    - If interpretability is crucial, use techniques like feature importance analysis or SHAP (SHapley Additive exPlanations) values to understand the factors contributing to churn predictions.

8. **Deployment:**
    - Once the model achieves satisfactory performance, deploy it in a production environment. This could involve integrating it into your CRM system, marketing platform, or other relevant systems.
    - Implement a mechanism to make real-time predictions based on new customer data.

9. **Monitoring and Maintenance:**
    - Continuously monitor the model's performance in the production environment.
    - Periodically retrain the model with updated data to adapt to changing customer behavior.
    - Address any concept drift (changes in data distribution) that may affect model accuracy over time.

10. Actionable Insights:

- Use the churn predictions to take proactive actions to retain customers.
- Develop strategies such as targeted marketing campaigns, personalized offers, or improved customer service for high-risk customers.

CONCLUSION

Customer churn prediction is a vital business strategy that uses data and machine learning to proactively identify and prevent customers from leaving. By collecting and processing relevant data, selecting suitable models, and implementing a well-structured process, businesses can make informed decisions to retain customers, boost revenue, and stay competitive. Continuous monitoring and adaptation are key to long-term success in customer retention.