Customer Churn Prediction Using Machine Learning

1. Problem statement

In the highly competitive energy market, a major gas and electricity utility company supplying SMEs faces increasing customer turnover as clients switch to alternative providers. This customer churn threatens market share and revenue stability. The company lacks tools and insights to proactively address churn, necessitating the identification of key indicators such as purchase frequency, customer complaints, and support ticket history. Leveraging these insights to implement targeted retention strategies and improve customer satisfaction is crucial for enhancing retention and ensuring sustained revenue growth.

2. Market/customer/business need assessment

a) Market Need

- Competitive Landscape: The energy market is characterized by a plethora of choices and competitive pricing, which makes it easy for customers to switch providers. This necessitates the implementation of advanced analytics to predict and mitigate churn.
- Regulatory Changes: Frequent changes in energy regulations require the company to stay agile and responsive, ensuring compliance while maintaining customer satisfaction.

b) Customer Need

- Personalized Experience: SMEs expect personalized services and targeted offers that cater to their specific needs and consumption patterns.

- Reliable Service: Consistent and reliable energy supply, along with prompt resolution of issues, is critical for customer retention.
- Transparency and Communication: Customers demand transparency in billing and regular communication regarding service updates and changes.

c) Business Need

- Data-Driven Insights: The company requires advanced tools to analyze customer data and identify churn indicators, enabling proactive retention efforts.
- Retention Strategies: Implementing effective retention strategies, such as personalized offers and discounts, is essential to reduce churn rates.
- Revenue Stability: By retaining customers and enhancing their lifetime value, the company can achieve sustained revenue growth and secure its market position.
- Customer Satisfaction: Improving overall customer satisfaction through data-driven approaches will foster loyalty and reduce attrition rates.

3. Target specifications and characteristics

The utility company's customer churn prediction model needs to meet certain goal criteria and attributes in order to address customer attrition in the competitive energy market. These goals will direct the model's creation and application, guaranteeing that it produces precise and useful insights:

I. **Model Accuracy**:

- **Prediction Accuracy**: The model should achieve a high level of accuracy (preferably above 90%) in predicting customer churn, minimizing false positives and negatives.
- **Precision and Recall**: Balancing precision and recall is essential to ensure the model identifies true churn risks without over- predicting.

II. **Data Requirements**:

- **Comprehensive Data Collection**: The model should utilize a wide range of customer data, including purchase frequency, billing history, customer complaints, support ticket history, and social media feedback.
- **Data Quality**: Ensuring high-quality, clean, and up-to-date data is crucial for the model's performance and reliability.

III. **Feature Engineering**:

- **Relevant Features**: Identify and incorporate key indicators of churn such as usage patterns, payment behaviour, customer interactions, and demographic information.
- **Feature Importance**: Assess the importance of each feature in predicting churn to prioritize the most impactful variables.

IV. Algorithm Selection:

- **Model Types**: Evaluate multiple machine learning algorithms, including logistic regression, decision trees, random forests, and gradient boosting, to determine the best-performing model.
- **Hyperparameter Tuning**: Optimize the chosen algorithm's hyperparameters to enhance model performance.

V. **Integration and Deployment**:

- **System Integration**: Ensure seamless integration with the company's existing customer relationship management (CRM) and billing systems for real-time churn prediction and intervention.
- **Scalability**: The model should be scalable to handle increasing data volumes as the company grows.

VI. **User Interface**:

- **Intuitive Dashboard**: Develop an intuitive dashboard for business users to easily access churn predictions, key indicators, and actionable insights.
- **Visualization Tools**: Include visualization tools to help users understand trends and patterns in churn data.

VII. **Actionable Insights**:

- **Retention Strategies**: Translate model predictions into actionable retention strategies, such as personalized offers, discounts, and engagement activities.
- **Customer Segmentation**: Segment customers based on churn risk to tailor retention efforts effectively.

VIII. **Performance Monitoring**:

- **Continuous Monitoring**: Implement continuous monitoring of the model's performance to ensure it remains accurate and relevant over time.

- **Periodic Retraining**: Periodically retrain the model with new data to maintain its predictive accuracy as market conditions and customer behavior evolve.

4. External searches

REFERENCES:

- Customer churning analysis using machine learning algorithms - ScienceDirect
- Customer - Churn Prediction Using Machine Learning | IEEE Conference Publication | IEEE Xplore
- Customer Churn Prediction Using Machine Learning - Javatpoint

DATASETS:

5. Benchmarking

Identify Existing Solutions

First, research the current customer churn prediction products or services available to SMEs in the energy market. These solutions can include:

- **Standalone Software**: Independent applications specifically designed for churn prediction.
- **Features by Energy Providers**: Tools and features integrated into services offered by energy providers.
- **Consulting Services**: Expert consulting firms offering churn prediction as part of their services.

Understanding these existing solutions provides a baseline for comparison and helps identify market gaps BCG can fill.

Analyze Features and Functionality

Next, compare the features and functionalities of existing solutions with BCG's proposed product. Key aspects to analyze include:

- **Data Integrations**: How well the product integrates with various data sources used by SMEs, such as billing systems and CRM platforms.
- **Churn Risk Scoring**: The ability to predict the likelihood of customer churn accurately.

- **Churn Reason Analysis**: Tools that help identify reasons why customers might churn.
- **Reporting Capabilities**: The quality and usability of reports generated by the tool.
- **User Interface**: The overall user experience, ensuring it's intuitive and easy for SMEs to use.

This comparison helps identify areas where BCG's product can excel or improve, focusing on delivering superior value to clients.

Evaluate Pricing Models

Analyzing the pricing models of competing solutions is crucial. Pricing models can include:

- **Subscription-Based**: Regular payments, typically monthly or annually, for continued access to the product.
- **Pay-Per-Use**: Charges based on the amount or frequency of use.
- **One-Time Fee**: A single upfront payment for perpetual access to the product.

Understanding these models helps BCG develop a competitive pricing strategy that is attractive to SMEs while ensuring profitability.

Assess Strengths and Weaknesses

By comparing the features, functionality, and pricing of BCG's product with those of competitors, BCG can identify its strengths and weaknesses. This assessment allows BCG to focus on unique selling points (USPs) and areas that require improvement. For example, if BCG's product offers superior data integration and user interface, these should be highlighted as key differentiators

6. Applicable Patents

Established Practice in Machine Learning

- **Machine Learning for Churn Prediction**: The use of machine learning algorithms for customer churn prediction is a well-established practice. While the core idea of using machine learning is not patentable, specific implementations or combinations of techniques might be patented.

- **Random Forests**: Random forests are a common and widely used machine learning algorithm. Patenting their general application for churn prediction would be difficult due to their prevalence.

Potential Patent Areas

- **Specific Techniques with Random Forests**: There could be patents for specific ways to use random forests for churn prediction. This includes unique feature engineering approaches, data pre-processing methods, or combining random forests with other algorithms.
- **Systems or Platforms for Churn Prediction**: Patents might cover comprehensive systems or platforms incorporating random forests for churn prediction. These patents could involve data integration tools, visualization dashboards, or specific churn intervention recommendations.

Recommendations

1.Conduct a Freedom-to-Operate (FTO) Search: An FTO search by a patent attorney can help identify relevant patents that might impact your product development. This search will provide a clearer picture of the patent landscape and potential risks.

2.Focus on Innovation: While random forests might be a core component of your product, focus on innovative aspects such as:

- **Unique Data Sources**: Use distinctive data sources that add value to the predictions.
- **Specific Churn Intervention Strategies**: Develop unique strategies for churn intervention based on the predictions.
- **User Interface Tailored for SMEs**: Create a user interface designed specifically for SMEs in the energy sector.

7. Applicable Constraint

Space

- **Minimal Physical Infrastructure**: Since this is a software product, physical space requirements are minimal. Cloud-based deployment can eliminate the need for physical infrastructure.

- **Data Storage**: SMEs may not generate vast amounts of data. The product should be designed for efficient data storage and processing, potentially leveraging cloud-based solutions for scalability and cost-efficiency.

Budget

- **Development Costs**: Building a robust machine learning model requires expertise and significant data resources. Balancing these development costs with affordability for SMEs is crucial.
    - **Pre-built Models**: Consider using pre-trained churn prediction models or open-source libraries to reduce development time and costs.
    - **Subscription Model**: Implement a subscription-based pricing model that scales with data usage or features accessed by SMEs. This makes the product more affordable and attractive to SMEs.

Expertise

- **Data Science & Machine Learning**: Developing and maintaining a churn prediction model requires specialized data science expertise.
    - **Collaboration**: Collaborate with universities or data science consultancies to access the necessary expertise and resources.
    - **Low-Code/No-Code Solutions**: Explore low-code/no-code platforms for model deployment if there is a lack of in-house data science resources. However, be aware that these platforms may limit customization and flexibility.

8. Business Model

Subscription Model (Tiered Pricing)

- **Free Tier**: A basic tier with limited functionality, such as a basic churn risk score, to attract new customers and showcase product value.
- **Paid Tiers**: Higher tiers offering advanced features like:
    - More granular churn risk scores
    - Detailed churn reason analysis
    - Historical trend analysis
    - Customized reporting and visualizations
    - Integration with intervention tools (e.g., email marketing platforms)

Pay-per-Use Model

- **Usage-Based Pricing**: Charge SMEs based on data usage or the number of churn predictions required, suitable for those with limited data or fluctuating prediction needs.

Freemium with Value-Added Services

- **Free Basic Version**: Offer core functionalities for free.
- **Paid Add-Ons**: Charge for additional services such as:
    - Expert consultations on churn reduction strategies
    - Implementation assistance for churn intervention plans
    - Custom data integrations

Freemium with Data Insights

- **Free Churn Prediction**: Provide basic churn prediction for free.
- **Paid Data Insights**: Charge for access to detailed data insights and visualizations, helping SMEs understand the reasons behind churn.

Partnership Model

- **Energy Provider Partnerships**: Collaborate with energy providers like PowerCo to offer the product as an add-on service to their SME customers. Revenue can be shared, or BCG can receive referral fees for each converted customer.

9. Concept Generation

Customer Churn (also known as customer attrition) refers to the phenomenon where customers cease their relationship with a company within a given time period. This can include cancelling subscriptions, ending contracts, or not making repeat purchases1. Here are the key steps for building an effective churn prediction model:

Define Churn Criteria and Goals:

Clearly define what constitutes churn for your business (e.g., cancelled subscriptions, inactive accounts).

Set specific goals for churn prediction, such as reducing churn rate or improving customer retention.

Data Collection and Preprocessing:

Gather relevant historical customer data (e.g., purchase frequency, engagement metrics, demographics).

Ensure data quality and integrity by handling missing values and outliers.

Select a Machine Learning Algorithm:

Choose an appropriate algorithm for churn prediction (e.g., logistic regression, decision trees, random forests, gradient boosting).

Split the data into training and testing sets, train the model, and optimize hyperparameters2.

Evaluate Model Performance:

Use testing data and cross-validation to assess how well the model predicts churn.

Adjust the model as needed to improve accuracy.

Remember that churn prediction is essential for businesses to retain customers, enhance overall performance, and address underlying issues that may lead to attrition.

10.Concept Development

Developing a concept for customer churn prediction involves several key steps. Here's a structured approach to guide you through the process:

1. **Define the Problem**
Objective: Understand why customers are leaving and identify those who are at risk of churning.
Scope: Determine the specific segment of customers to focus on (e.g., by product line, region, etc.).

2. **Understand the Business Context**
Customer Lifecycle: Map out the typical customer journey.
Churn Definition: Clearly define what constitutes churn for your business (e.g., cancellation of service, no purchase within a certain period).

3. **Data Collection**

Internal Data Sources: CRM systems, sales data, customer service records, usage logs, etc.
External Data Sources: Market trends, social media, economic indicators.
Data Points: Transaction history, interaction history, demographics, customer service interactions, etc.

4. **Data Preparation**
Data Cleaning: Remove duplicates, handle missing values, standardize data formats.
Feature Engineering: Create relevant features such as frequency of use, recency of last interaction, customer lifetime value, etc.
Normalization: Scale features to ensure consistency across different data types.

5. **Exploratory Data Analysis (EDA)**
Descriptive Statistics: Summarize the main characteristics of the data.
Visualizations: Use charts and graphs to understand data distributions and relationships.
Correlation Analysis: Identify relationships between features and churn.

6. **Model Selection**
Algorithms: Choose from various machine learning algorithms such as Logistic Regression, Decision Trees, Random Forests, Gradient Boosting Machines, Neural Networks, etc.
Evaluation Metrics: Select appropriate metrics like accuracy, precision, recall, F1-score, ROC-AUC, etc.

7. **Model Training and Validation**
Training Data: Split data into training and testing sets.
Cross-Validation: Use techniques like k-fold cross-validation to ensure model robustness.
Hyperparameter Tuning: Optimize model parameters for better performance.

8. **Model Deployment**
Integration: Deploy the model within the existing customer relationship management system.
Real-Time Predictions: Ensure the model can make real-time predictions as new data comes in.

9. **Monitoring and Maintenance**
Performance Tracking: Continuously monitor the model's performance over time.
Feedback Loop: Use new data to retrain and improve the model.
Model Updates: Regularly update the model to adapt to changing customer behaviour and market conditions.

10. **Actionable Insights and Strategy**
Churn Prevention: Develop strategies to retain at-risk customers based on insights from the model (e.g., personalized offers, targeted marketing campaigns, improved customer service).
Business Decisions: Use predictions to inform broader business decisions and strategies.

11. **Ethical Considerations**
Privacy: Ensure compliance with data privacy regulations.
Bias: Regularly check the model for biases and ensure fairness in predictions.

This framework provides a comprehensive approach to developing a customer churn prediction model. Tailoring each step to the specific needs and context of your business will help create a more effective and accurate model.

11. Final Product Prototype/ Product Details

Feasibility:

1. **Business Relevance**

Impact: High potential to reduce customer churn, increase retention, and boost revenue.

Stakeholder Support: Typically garners strong support from marketing, sales, and customer service teams.

2. **Data Availability**

Internal Data: Customer transaction history, interaction logs, demographics, usage patterns.

External Data: Market trends, competitor data, economic indicators.

Requirement: Sufficient historical data for model training and validation.

3. **Technical Capability**

Tools: Availability of machine learning tools and platforms (e.g., Python, R, cloud services).

Skills: Data science expertise within the team or ability to hire/contract the necessary skills.

4. **Implementation**

Integration: Feasible integration with existing CRM systems for real-time prediction.

Scalability: Ability to handle large volumes of data and provide predictions at scale.

5. **Cost**

Initial Investment: Costs for data collection, storage, processing, and initial model development.

Maintenance: Ongoing costs for model monitoring, updates, and retraining.

6. **Ethical and Legal Considerations**

Privacy: Compliance with data privacy laws (e.g., GDPR, CCPA).

Bias: Ensuring model fairness and avoiding discrimination.

Customer churn prediction is generally feasible with the right data, technical capabilities, and stakeholder support. It offers significant business benefits by enabling proactive customer retention strategies.

Viability:

1. **Strategic Value**

Retention Impact: High potential to improve customer retention and reduce churn.

Revenue Growth: Opportunity to boost revenue by retaining more customers.

2. **Data Availability and Quality**

Sufficient Data: Must have historical customer data (transactions, interactions, demographics).

Data Quality: Clean, complete, and relevant data is crucial for accurate predictions.

3. **Technical Infrastructure**

Tools and Platforms: Access to machine learning tools (e.g., Python, R) and platforms (cloud services) is essential.

Expertise: Availability of data science skills within the team or through external resources.

4. **Implementation Feasibility**

Integration: Can be integrated into existing CRM systems for real-time use.

Scalability: Capable of handling large datasets and generating real-time predictions.

5. **Cost and Resources**

Initial Investment: Costs for data collection, storage, processing, and model development.

Ongoing Maintenance: Resources needed for model updates, monitoring, and retraining.

6. **Ethical and Regulatory Compliance**

Data Privacy: Adherence to data privacy regulations (e.g., GDPR, CCPA).

Bias and Fairness: Ensuring the model does not introduce or perpetuate bias.

Customer churn prediction is viable if the organization has the necessary data, technical capabilities, and resources. It offers significant strategic value by enabling proactive retention efforts and improving overall customer satisfaction.

Monetization:

1. Revenue Retention

- Reduced Churn**:** By identifying at-risk customers, companies can implement targeted retention strategies, thereby reducing churn rates and preserving revenue.
- Customer Lifetime Value**:** Increased retention leads to higher customer lifetime value (CLV).

2. Targeted Marketing

- Personalized Offers: Tailoring promotions and offers to at-risk customers can increase the likelihood of retention and additional purchases.
- Cost Efficiency: More effective allocation of marketing budgets by focusing on customers who are most likely to churn.

3. Customer Experience

- Proactive Support: Offering proactive customer service interventions to at-risk customers can improve satisfaction and loyalty.
- Product Improvement: Insights from churn prediction can inform product or service enhancements, leading to better customer retention.

4. Strategic Planning

- Forecasting: Better revenue forecasting and planning based on predicted churn rates.
- Resource Allocation: More efficient allocation of resources to high-risk areas, improving overall business performance.

5. Competitive Advantage

- Market Position: Maintaining a lower churn rate compared to competitors can enhance market position and brand loyalty.
- Customer Insights: Gaining deeper insights into customer behaviour and preferences helps in staying ahead of competitors.

Monetizing customer churn prediction involves leveraging insights to retain more customers, enhance marketing efficiency, improve customer experience, and make informed strategic decisions. This results in increased revenue, higher customer lifetime value, and a stronger competitive position.

Step 4: Financial Modelling : Here, we did an analysis based on the churning of the customers for every 3 months of their satisfaction.

Total Customers = 14,606

Churn Rate (Found by using Random Forest model) = 10%

Average Revenue per Customer = $180

CLV = $2160

Customer Acquisition Cost = $18

Retention Cost Per Customer = $9

Predicted Churn Customers = 1460

Revenue at Risk = $262800

Total Retention

Cost = $13140

Net Revenue Impact = $249660

It shows an Exponential Market Trend:

Initial Profit: CLV (Customer Lifetime Value) =$2,160

Growth Rate: Churn rate = 10% (in decimal 0.1)

Exponential Model Equation: y = 2160 * (exp (0.1*t))

Conclusion

Customer churn prediction is a strategic tool that can significantly enhance a company's ability to retain customers and improve overall business performance. By leveraging historical and real-time data, businesses can identify customers at risk of leaving and implement targeted retention strategies. Overall, customer churn prediction is a viable and valuable approach for businesses looking to enhance customer retention, increase revenue, and gain competitive advantage. With the right data and resources, it can provide actionable insights and significant business benefits.