Subject: Hypothesis Testing and Predictive Modelling for Churn Reduction

Dear Sir/Madam

I hope this message finds you well. I wanted to share some initial thoughts and steps on how we can address the significant churn problem faced by our client, a major gas and electricity utility.

Hypothesis Statement: Our hypothesis is that customer churn is primarily driven by price sensitivity.

Major Steps to Test the Hypothesis:-

1. **Data Collection**: To test this hypothesis effectively, we would require historical customer data from our client. This data should include customer demographics, usage patterns, billing information, contract details, historical churn data, and any other relevant features. Additionally, external data sources, such as market price fluctuations, could be valuable.

2. **Data Preprocessing**: The collected data needs to be cleaned and preprocessed. This involves handling missing values, encoding categorical variables, and scaling numerical features. Special attention should be given to feature engineering to extract insights related to price sensitivity, such as historical price changes

3. **Exploratory Data Analysis (EDA):** EDA is crucial for gaining initial insights into the data. We should visualize and analyse customer churn patterns, identify correlations, and assess the impact of pricing-related features on churn.

4. **Model Selection:** We would employ predictive modelling techniques to build a churn prediction model. Ensemble methods like Random Forest or Gradient Boosting, as well as logistic regression, are common choices. The target variable would be churn (binary: yes/no).

5. **Feature Importance:** We'll determine which features have the most significant impact on churn prediction. This will help us identify which price-related variables are influential.

6. **Model Evaluation:** We'll evaluate the model's performance using metrics such as accuracy, precision, recall, F1-score, and ROC-AUC. We may consider cross-validation to ensure the model's generalizability.

7. **Discounting Strategy:** Once the model is built and validated, we can use it to predict which SME customers are at high risk of churning. The strategy would involve offering a 20% discount to these high-risk customers and monitoring their response over a specific period.

8. **A/B Testing:** To assess the effectiveness of the discounting strategy, we can conduct A/B testing. We would randomly assign customers to two groups: one that receives the discount offer and another that does not. By comparing churn rates between the two groups, we can determine if the discount strategy has a statistically significant impact on reducing churn.

9. **Continuous Monitoring:** After implementing the strategy, it's crucial to continuously monitor customer behaviour, including churn rates and responses to discounts. This will help us make data-driven adjustments as needed.

Data Requirements from the Client:-

- Historical customer data (demographics, usage, billing, contracts, churn history).
- Price-related data (pricing changes, market fluctuations).
- External market data (if available).
- Access to customer interaction records (e.g., customer service calls, emails).

I believe that by rigorously testing our hypothesis using predictive modelling and a discounting strategy, we can help our client mitigate churn effectively in the SME segment.

Please let me know if you have any further insights or specific client requests that we should consider during this process. I look forward to discussing this in more detail.

Best Regards,
Amey Desai