Subject: Testing the Hypothesis of Price Sensitivity in Customer Churn

Dear Associate Director,

I hope this email finds you well. I've carefully considered the hypothesis that customer churn in the SME segment is driven by price sensitivity. Here are my initial thoughts on how to approach this problem and test this hypothesis effectively.

Formulating the Hypothesis as a Data Science Problem:

To formally test the hypothesis that churn is driven by price sensitivity, we need to frame it as a predictive modeling problem. Specifically, we can define the problem as follows:

Hypothesis: Customer churn in the SME segment is influenced by price sensitivity.

Data Science Problem: Build a predictive model that can identify SME customers at risk of churning based on their price sensitivity.

Major Steps to Test the Hypothesis:

1. Data Collection:

- We need access to historical customer data, including customer profiles, usage patterns, billing information, and churn status.

- We also require data on price changes and discount histories, which should be provided by the client.

2. Data Preprocessing:

- Clean the data to handle missing values and outliers.

- Create a target variable, such as a binary indicator of churn (1 for churned, 0 for retained).

- Feature engineering: Extract relevant features related to price sensitivity, such as customer demographics, historical price changes, contract duration, and usage behavior.

3. Model Selection:

- Choose an appropriate machine learning model for prediction. This could be a classification model, as we are predicting churn (e.g., Logistic Regression, Random Forest, Gradient Boosting).

- Train the model on historical data with churn labels.

4. Model Evaluation:

- Assess the model's performance using relevant metrics such as accuracy, precision, recall, and F1-score.

- Use techniques like cross-validation to ensure the model's generalizability.

5. Feature Importance Analysis:

- Analyze the importance of different features, particularly focusing on those related to price sensitivity. This will help us understand which factors are driving churn.

6. Model Deployment:

- Implement the predictive model into the client's system to use it on the 1st working day of every month.

7. Post-Deployment Analysis:

- Continuously monitor the model's performance and retrain it with new data as needed.

- Regularly assess the impact of offering a 20% discount to customers identified as at risk of churning.

8. A/B Testing (Optional):

- To evaluate the effectiveness of the 20% discount, consider running A/B tests. Offer the discount to a random subset of at-risk customers and compare their churn rates to those who were not offered the discount.

Data Requirements:

To proceed with this analysis, we will require the following data from the client:

- Historical customer data, including churn status.

- Billing and usage data.

- Price change history.

- Information on any past discounts offered to customers.

Next Steps:

Once we have access to the necessary data, I propose that we start by conducting exploratory data analysis (EDA) to gain insights into the data and understand the behavior of SME customers. This will help us refine our feature selection and engineering process.

We can then proceed with the development of the predictive model and test its performance.

Please let me know if you have any further questions or if you would like to discuss this approach in more detail. I'm eager to start working on this hypothesis and help our client mitigate churn in the SME segment.

Best regards,

Saad Rasheed