

Subject : Proposal for Testing the Hypothesis on Churn Driven by Price Sensitivity

Dear Sir/Mam,

I hope this email finds you well. Following our team meeting on hypotheses surrounding SME

customer churn, I've had the opportunity to delve deeper into the hypothesis that customer churn is

driven by price sensitivity. Below, I outline the steps and considerations for testing this hypothesis.

Formulating the Hypothesis as a Data Science Problem:

The hypothesis we are testing can be framed as a predictive modeling problem.

Specifically, we aim to

predict customer churn based on various features, with a focus on price sensitivity. The goal is to

understand which customers are more likely to churn at their current price and, subsequently, identify

those who could be retained with a 20% discount.

Steps to Test the Hypothesis:

1. Data Collection: Gather historical data on customer behavior, including churn events.

Collect

information on pricing changes and historical discounts offered to customers.

2. Feature Engineering: Identify relevant features that may influence churn, including: Price changes

and sensitivity, Contract duration, Usage Patterns, Demographic information and Customer tenure.

3. Data Pre processing: Handle missing values and outliers. Normalize or scale numerical features.

Encode categorical variables appropriately.

4. Model Selection: Consider using machine learning algorithms such as logistic regression, decision

trees, or ensemble methods. Evaluate model performance using metrics like accuracy, precision,

recall, and F1-score.

5. Model Training and Validation: Split the dataset into training and validation sets. Train the model

on historical data and validate its performance on a separate dataset.

6. Interpretability Analysis: Assess the importance of each feature, with a focus on price sensitivity.

Understand how changes in price affect the likelihood of churn.

7. Prediction for Monthly Discount Offers: Apply the trained model to predict which customers are at a higher risk of churning. Recommend a 20% discount to customers identified as high-risk.

8. Monitoring and Iteration: Regularly monitor the model's performance and update it as new data

becomes available. Iterate and refine the model based on ongoing insights and feedback.

Data Requirements:

To proceed with the analysis, we would need access to the following data from the client:

- Historical customer data, including churn events.
- Pricing information and changes over time.
- Details on discounts offered to customers.
- Any additional relevant data such as customer demographics.

Timeline :

Considering the client's plan to use the predictive model on the 1st working day of every month, we

should aim to have a working model well in advance. A tentative timeline could include data

collection and preprocessing in the first two weeks, model development and validation in the third

week, and the final model ready for deployment by the end of the month.

I look forward to discussing this proposal further and welcome any additional insights or suggestions

from the team.

Best regards,

Lachiket Warule

Data Science Intern