Subject: Testing the Hypothesis: Churn Driven by Price Sensitivities

Dear Associate Director,

I hope this email finds you well. As per our discussion in the team meeting, I have compiled my thoughts on how we should test the hypothesis that churn in the SME segment is driven by customers' price sensitivities. I believe that conducting a comprehensive analysis and building a predictive model will provide valuable insights for our client, PowerCo. Here's an outline of the steps we can take:

Hypothesis: Churn in the SME segment is driven by customers' price sensitivities.

Data Science Problem: Predicting customer churn in the SME segment based on price sensitivities.

Major Steps to Test the Hypothesis:

1. Data Collection and Preparation:

- Gather historical data on customer churn, pricing, customer demographics, usage patterns, and any other relevant information.
- Identify the target variable, which is churn (binary: churned or not churned) within a specific time period.
- Extract and pre-process features such as customer tenure, billing amount, contract length, customer type, and any other potential indicators of price sensitivity.

2. Exploratory Data Analysis (EDA):

- Perform descriptive analysis to understand the distribution of churn, pricing, and other relevant variables.
- Conduct correlation analysis to identify relationships between variables and churn, with a focus on price-related features.
- Segment the customer base based on characteristics like industry, location, or usage to identify any variations in price sensitivity within subgroups.

3. Model Development:

- Select an appropriate machine learning algorithm(s) for predicting churn based on price sensitivity. Options include logistic regression, decision trees, random forests, or gradient boosting machines.
- Split the historical dataset into training and testing sets for model development and evaluation.
- Utilize techniques such as feature selection (e.g., feature importance, recursive feature elimination) to identify the most relevant features for prediction.

4. Model Training and Evaluation:

- Train the predictive model using the training set and optimize hyperparameters using techniques like cross-validation.
- Evaluate model performance using appropriate evaluation metrics such as accuracy, precision, recall, and F1-score.
- Consider using additional techniques like ROC curves and AUC to measure model performance.

5. Predictive Model Application:

- Apply the trained model on the 1st working day of each month to generate churn predictions for the upcoming month.
- Identify customers who are most likely to churn at their current price point based on the churn predictions.
- Determine the appropriate discount amount required to incentivize at-risk customers to stay, considering factors such as churn probability and potential revenue loss.

6. Monitoring and Iteration:

- Continuously monitor the performance of the predictive model to ensure its accuracy and reliability.
- Refine hypotheses and insights based on the findings from the predictive model.
- Iterate on the model and methodology as necessary to improve its predictive power and usefulness.

By following these steps, we can systematically test the hypothesis that churn in the SME segment is driven by customers' price sensitivities. The predictive model will enable us to identify customers at risk and determine the optimal discount amount to discourage churn effectively.

Thank You

Regards

Himanshu Kumar

Data Science and Analytics – Virtual Intern