Scotia Data Science Day
Winter 2024

Background

In today's digital age, customer feedback serves as a crucial barometer for the success and user satisfaction of mobile and online applications, particularly for financial institutions. Analyzing customer app review is integral to enhancing the customer experience. These reviews offer a wealth of unstructured feedback directly from the end users, providing invaluable insights into customer satisfaction, usability issues, and desired features. By meticulously examining these comments, financial institutions can identify recurring issues, understand customer sentiment, and gauge the effectiveness of new features. Furthermore, this analysis not only helps in fine-tuning the app's features but also demonstrates a commitment to listening to customer needs, thereby fostering loyalty and trust. This feedback loop is vital in a competitive market where user experience can significantly influence customer retention and acquisition. In the competitive financial sector, where digital convenience and reliability are paramount, leveraging such customer feedback is essential for staying ahead and ensuring a seamless, positive user experience.

Scenario

As a young and talented data scientist, you've recently joined the customer advocacy team, collaborating with like-minded analysts. Your team's primary goal is to analyze customer feedback to enhance the customer experience.

In reviewing Scotiabank's mobile app customer reviews, you noted datasets containing irrelevant reviews from external sources. Your team's objective is to segregate and eliminate these irrelevant "Easter egg" reviews, then focus on detecting popularity of 20 selected topics to discover the relevant reviews. Moreover, you're responsible for deriving insights and formulating recommendations to improve customer experience based on your data analysis. This will require you to analyze the data, identify patterns and trends, and develop business strategies to improve customer experience.

Ask

Your team needs to use advanced analytics/AI approach to analyze customer app review to improve customer experience based on the given available data. Furthermore, you need to look for any insights from the data that can help you conduct business analysis to answer the following questions:

- *What are frequent and popular topics among Scotiabank mobile app review, can you identify the popularity of 20 topics among given topics from data dictionary, how did you arrive at the conclusion?*
- *What are some of the customers' needs for Scotia mobile app? desired features, pain points?*
- *If you are looking to build/fix one feature in Scotiabank mobile app to improve customer experience what will that be?*
- *Are there some external sources of customer reviews mixed with the data? Can you identify those "Easter egg" reviews and what are those reviews about?*
- *Can you make any long-term suggestions to improve customer experience?*
- *How did you arrive at the conclusions?*

Tools and Methods

You can use any common statistical or programming languages such as Python, R, etc. as well as open-source tool e.g. Chatgpt. You can try different AI/ machine learning methods like natural language processing, large language model, generate AI etc.

Teams

Each team should have up to five students. Students who have registered without a team will be grouped randomly into teams.

Data

- 9175 records in the dataset
- Each record is a customer app review, one ID column (`REVIEW_ID`) to uniquely identify a customer review
- Description of column (sheet one) and potential topics (sheet two) are available in data dictionary

Submission

Submission should be made to the private MS Teams channel before 12 pm noon on Friday, Feb 2, 2024. Participants are expected to submit all four items below and follow the naming conventions. Details of what each item should include are provided in the Judging Criteria document.

1) The detection result should be submitted with 2 separated files in csv format
   - One csv file named `[team_name]_topics.csv` with three columns:
     - `TOPIC_NAME` (exact name with no space from data dictionary),
     - `RANKING` (value of 1-20, 1 indicating highest ranking topic),
     - `NUMBER_OF_OCCURANCE` (value of 0 – 9999, number of reviews related to the topic)
   - One file with one column: `Review_ID` (reviews that are classified as an "Easter egg" by your team) `[team_name]_easter_eggs.csv`
2) Slideshow presentation (Maximum 7 slides)
   - `[team_name]_slides.pptx/pdf`
3) Brief write-up (Maximum 500 words)
   - `[team_name]_brief.pdf`
4) Code in working state (in a zip file)
   - `[team_name]_code.zip`

Assessment

There are two rounds of assessment evaluated by Scotiabank Senior Management and University of Waterloo faculty members. Detailed judging criteria are provided in the Judging Criteria document.

Criteria for Round 1:

- Performance of the topic occurrences and "Easter eggs" review detection based on dataset using following measurement: topic popularity measured by number of occurrences review contains the selected topic , as well as # of correct "Easter egg" reviews identified (wrongly identified review will receive penalty)
- Insights, recommendation, and visualization in slides
- Evidence of analytical rigor and creativity
- Use of modern LLM techniques or advanced analytics approach to derive business insight

The six best teams (determined by overall score calculated based on Judging Criteria) will be selected to present their work in a 7-minute presentation to the judges + 3-minute Q&A

Criteria for Round 2:

- Clarity and organization of thought
- Overall presentation
- Analytical insight

Prize Structure

- First Place: $2,000
- Second Place: $1,500 x 2
- Third Place: $1,000 x 3
- Invite First Place to meet Scotiabank CID&A Teams / Execs for snack/dinner on Friday

Important Notes

- Use MS Teams common channel to drop your logistics questions or attend the mentorship session on Monday, Wednesday, Thursday for technical questions
- Formal Kick-off Friday (Jan 26th), please join to check-in and ask questions!
- Please submit your team result with required format on time & prepare presentation ahead of time
- Final presentations on Friday (Feb 2nd), team must attend in person to qualify for Top 6 winning teams

Thank you for your interest, and good luck!