Amazon Sales Chatbot

Phase 1: Proposal Report

Presented by

- Minn Banya
- Wut Yee Aung
- Kyi Thin Nu
- Thamakorn Kiattikaikul
- Noppawee Teeraratchanon

Natural Language Processing

Recall

Background: The Rise of Sales Chatbots

Sales chatbots, specifically tailored for Amazon, are AI-driven virtual assistants designed to engage customers in conversational interactions, guide them through the Amazon sales process, and facilitate purchases.

Key Factors Driving Adoption:

- 24/7 Availability
- Scalability:
- Enhanced Customer Experience.
- Cost-Efficiency
- Data Insights

Sales chatbots tailored for Amazon offer sellers a modern and effective solution for enhancing customer engagement, driving sales, and gaining valuable insights into Amazon customer preferences and behavior.

Problem and Motivation

Problem

- Many customers might abandon the sales process due to lack of immediate assistance or the inability to get answers to their questions quickly.
- The problem also lies in the fact that customers often require accurate information and may have follow-up questions that a pre-programmed chatbot cannot adequately address. This leads to frustration and potential loss of sales as customers seek answers elsewhere or abandon their purchase altogether. Traditional chatbots, lacking the ability to adapt to nuanced inquiries, struggle to provide satisfactory responses in such scenarios.

Problem and Motivation

Motivation

- Motivation arises from the pressing need to bridge this gap in customer service to ensure a seamless and satisfying shopping experience. By addressing this challenge, businesses can enhance customer trust, loyalty, and ultimately drive sales growth on platforms like Amazon.
- Motivation stems from the desire to increase sales, improve customer satisfaction, and reduce the workload on human sales representatives.

Solution Requirements

- Natural Language Processing (NLP): Ability to understand and respond to customer inquiries in natural language.
- Faster and more accurate information retrieval
- Preprocessing the database by separating into labeled bins/ vector stores
- Classifier to determine which vector store the LLM should retrieve information from based on user input
- Designing a suitable prompt template
- Fine-tuning the LLM on Amazon QA data
- Experiments carried out to determine the best model for both classifier and LLM
- Adding additional components to the models, such as Active RAG
- Ablation study to determine which components have the most impact
- Web application that can serve as a chat window for the customer

Architecture of the Solution

- Data Source and Preprocessing
- NLP
  - Classification Part
    - MLflow will be used to save the models and see the performance of the models (select the best performance model as our best model)
  - LLM Part
    - Database will be split into vector stores
    - MLflow will be used for experiment logging and model integration
    - Fine-tuning tested and included if there is a positive impact on the performance
- Design test cases and evaluations for the system's performance
- Web Design
  - Flask as backend framework to handle interaction between user requests and model responses.
  - HTML as frontend framework for the user interface.
  - MLflow for managing machine learning models including experiment tracking, and model versioning.
  - GitHub will manage codebase changes while also tracking modifications in application code changes.

Experiments

We will do the following experiments:

- User Testing: Conduct usability testing with a sample group of potential customers to evaluate the chatbot's effectiveness and user experience.
- A/B Testing: Compare the performance of the chatbot with and without certain features or algorithms to measure their impact on sales conversion rates.
- Data Analysis: Analyze the data collected from user interactions to identify patterns and areas for improvement.

Task Distribution

- Data Collection - Thamakorn and Kyi
- Modeling
  - Classification - Noppawee
  - Language Model - Minn
- Web Application - Wut Yee
- Experiment Testing - Thamakorn and Kyi
- Reporting and Others - Thamkorn and Kyi

Our Project GitHub Repository

https://github.com/minnbanya/NLP-Project

Our Data Source

Dataset information

- Amazon product dataset (small) - https://www.kaggle.com/datasets/promptcloud/amazon-product-dataset-2020
- Amazon product dataset (large) - https://www.kaggle.com/datasets/piyushjain16/amazon-product-data/data
- Amazon QA dataset - https://cseweb.ucsd.edu/~jmcauley/datasets/amazon/qa/

Thank you
For your attention