Development of FAQ Chatbot using BERT Embeddings and FAISS

Raushan Kumar

May 19, 2024

1 Introduction

In this report, we present the development of a Frequently Asked Questions (FAQ) chatbot leveraging state-of-the-art natural language processing techniques. The chatbot utilizes BERT (Bidirectional Encoder Representations from Transformers) embeddings for semantic similarity matching and FAISS (Facebook AI Similarity Search) for efficient nearest neighbor search.

2 Problem Statement

The objective is to develop a chatbot capable of providing relevant answers to user queries by leveraging a pre-existing FAQ dataset.

3 Approach

3.1 Data Acquisition and Preprocessing

- The FAQ data is obtained from a CSV file containing questions and corresponding answers.
- The data is loaded into memory using the pandas library.

3.2 BERT Model Integration

- We integrate a pre-trained BERT model and tokenizer from the Hugging Face transformers library.
- BERT embeddings are generated for both the FAQ questions and user queries.

3.3 FAISS Indexing

- We employ the FAISS library to create an index of FAQ embeddings for efficient similarity search.
- The FAISS index enables fast retrieval of the most relevant FAQ based on the similarity of embeddings.

3.4 Chatbot Implementation

- The chatbot interacts with users through a command-line interface.
- It continuously prompts the user for input and provides appropriate responses based on the most relevant FAQ answer.
- The chatbot maintains a conversation history to provide contextually relevant answers.

4 Results

The developed FAQ chatbot demonstrates effective performance in retrieving relevant answers to user queries. The integration of BERT embeddings and FAISS indexing ensures accurate and efficient matching of user queries with FAQ entries.

5 Conclusion

In conclusion, the developed FAQ chatbot represents a successful implementation of advanced natural language processing techniques for information retrieval tasks. By leveraging BERT embeddings and FAISS indexing, the chatbot provides users with timely and accurate responses to their queries, enhancing user experience and efficiency in accessing FAQ information.

6 Future Work

- Integration of additional features such as sentiment analysis and named entity recognition to enhance chatbot capabilities.
- Deployment of the chatbot as a web service or integration with existing messaging platforms for wider accessibility.
- Continuous improvement and fine-tuning of the chatbot's performance based on user feedback and usage analytics.