StonyBrook AI ChatBot

Rishith Kyatham

Introduction

Introduction to Stony Brook University Chatbot The Stony Brook University Chatbot is an intelligent, interactive web application designed to assist users with inquiries related to Stony Brook University. Built using Streamlit, a popular framework for creating interactive web applications in Python, this chatbot leverages advanced natural language processing (NLP) capabilities to provide accurate and relevant information. By integrating various components such as web scraping, data storage, and conversational AI, the chatbot offers a seamless user experience for accessing university-related information.

Key Features

User-Friendly Interface: The chatbot features a clean and intuitive interface built with Streamlit, enabling users to interact easily by typing questions and receiving responses in real-time.
Dynamic Content Retrieval: Utilizing web scraping techniques, the chatbot extracts and processes information from Stony Brook University's official website and academic bulletins, ensuring that the responses are based on the latest available data.
Advanced NLP Integration: Powered by OpenAI's GPT-3.5-turbo model, the chatbot can understand and respond to a wide range of questions, providing detailed and contextually accurate answers.
Chroma Vector Store: The chatbot employs a Chroma vector store to manage and retrieve document embeddings efficiently. This vector store ensures quick access to relevant information, enhancing the response time and accuracy.
Automated Data Handling: The chatbot includes functionalities for automatic data fetching, processing, and storage. The data file is periodically refreshed to maintain up-to-date information, and documents are stored in chunks for efficient retrieval.

Components and Workflow

Streamlit Application (main.py): The core of the application, main.py, sets up the Streamlit interface. It includes user input fields, buttons, and displays for interaction. The app also handles the initial setup, including data file management and document storage.
Helper Functions (helper_functions.py): This module contains essential functions for managing the vector store, storing documents, and generating responses. It defines how documents are fetched, processed, and stored in the vector store, as well as how responses are generated using the Conversational Retrieval Chain.
Web Crawler (web_crawler.py): The web crawler module retrieves and processes text content from specified URLs. It uses BeautifulSoup to parse HTML and extract meaningful text, which is then converted into document chunks for storage.
Environment Configuration: The chatbot securely loads API keys and other sensitive information from an environment file (env.sh), ensuring that credentials are not hard-coded into the source code.

Implementation Details

Data File Management: Upon startup, the existing data file (vector_store.json) is deleted to avoid redundancy. New documents are then fetched and stored to ensure the vector store contains the latest information.
Document Storage: Using the store_docs function, the chatbot retrieves and processes webpage content from Stony Brook University's website. The processed documents are stored in a Chroma vector store, which is then used to provide relevant responses.
Conversational Model: The get_response function leverages a Conversational Retrieval Chain that integrates the OpenAI GPT-3.5-turbo model and the Chroma vector store. This combination allows the chatbot to retrieve relevant information efficiently and generate accurate answers.

Conclusion

The Stony Brook University Chatbot is a sophisticated tool designed to facilitate access to university information. By integrating web scraping, vector storage, and advanced NLP models, it provides a reliable and interactive platform for users to obtain timely and accurate responses to their queries. This chatbot represents a significant step towards leveraging AI to enhance user engagement and information dissemination at Stony Brook University.

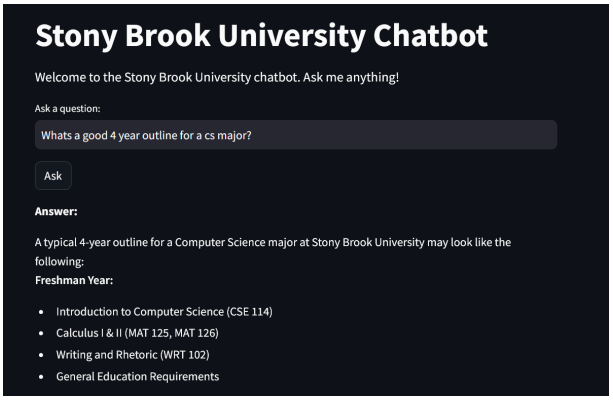

Fig. 1. StonyBrook ChatBot