Project Documentation

PubMed Article Summarizer Using GPT-3.5

By: Abdul Sami Khan, BSDS, 22i-2062

Introduction

This project involves the development of a web application that summarizes PubMed articles using OpenAI's GPT-3.5 model. The application is built using Flask and integrates functionalities for storing and reading summaries from a text file. The purpose of this project is to demonstrate the ability to summarize scientific articles and display both the original text and its summary on a web interface.

Background

I have been working on chatbots since last year. Previously, I developed a financial chatbot assistant with real-time data access using a Flask application and a file management system with user authentication. This experience laid the foundation for my current project, where I leveraged my knowledge and codebase to create a PubMed article summarizer.

Learning Journey

I learned Flask during my Big Data Lab in the 4th semester. This project is a culmination of my continuous efforts to enhance my skills and improve my projects based on new knowledge and techniques I acquire.

Project Structure

The project consists of the following components:

1. **Flask Application**: The core of the web application.
2. **GPT-3.5 Summarization**: Integration with OpenAI's GPT-3.5 for generating summaries.
3. **File Management System**: Storing and reading summaries from a text file.
4. **Web Interface**: User interface for inputting articles and displaying summaries.

Implementation

1. Flask Application

The Flask application serves as the backend for the web interface. It handles routes, form submissions, and integrates with the GPT-3.5 API for summarization.

2. GPT-3.5 Summarization

The summarization functionality is achieved using OpenAI's GPT-3.5 model. The model is called via the OpenAI API to generate summaries for the provided articles.

3. File Management System

Summaries are stored in a text file, and a function is provided to read and display these summaries. This ensures that users can view previously summarized articles.

4. Web Interface

The web interface is built using HTML and allows users to:
- Select articles from different datasets.
- Submit articles for summarization.
- View original articles and their summaries.
- Access a page displaying all stored summaries.