Preliminary Analysis of Datasets using GPT and Wolfram Language

Overview

This project aims to facilitate preliminary analysis of datasets using language models (GPT). It provides a web-based interface for users to upload datasets in various formats and pose questions about the data, receiving insights and visualizations generated by the models.

Files Included

- **app.py:** Initial version of the application.
- **main.py:** Final version with all functionalities and models integrated.
- **workingapp.py:** Contains additional methods and functionalities experimented during development.
- **templates:** Directory containing HTML templates and CSS files for web app styling.
- **requirements.txt**: List of all required Python packages and their versions.

How to Run

To run the application, follow these steps:

1. Clone or download the project repository to your local machine.
2. Navigate to the project directory in your terminal or command prompt.
3. Install the required dependencies by running:

```
pip install -r requirements.txt
```

4. Run the project using the following command:

```
streamlit run main.py
```

5. The application will automatically open in your default web browser.

Note: This web application is complete, and all the models can be accessed if locally installed.

Usage

- Upload your dataset in CSV, Excel, PDF, or JSON format.
- Pose questions about the dataset using the provided text input field.
- Click the "Analyze Data" button to generate insights and visualizations.
- Explore the generated results and visualizations to gain preliminary insights into your dataset.

Dependencies

The project requires the following dependencies:

```
streamlit==1.32.2
pandas==1.4.3
numpy==1.22.2
plotly==5.3.1
seaborn==0.11.2
fitz==0.0.1.dev2
langchain==0.1.13
langchain-community==0.0.29
langchain-experimental==0.0.55
langchain-openai==0.1.3
.
.
.
```

Install these dependencies using '*pip install -r requirements.txt*'.