

Infinite Interact: Multi-Agent Operational Suite

CMPE 258: Project Report

Shawn Chumbar

Aagam Shah

Sajal Agarwal

Dhruval Shah

Abstract

The Infinite Interact: Multi-Agent Operational Suite is a comprehensive project designed for CMPE 258, aimed at leveraging the power of Large Language Models (LLMs) in conjunction with LangChain and OpenAI to create a suite of AI agents. These agents are capable of performing a variety of tasks, ranging from video transcription and CSV querying to presentation generation, README file creation, and web scraping. Our team, Bay Area Rockers, developed five distinct agents, each tailored to address specific user needs. This report details the development process, functionalities, and applications of these agents, showcasing their potential to streamline and automate diverse operational tasks.

Introduction

The advent of Artificial Intelligence (AI) has revolutionized the way we interact with and utilize data. In the realm of software development and data analysis, AI-driven tools have become indispensable for automating repetitive tasks, enhancing productivity, and extracting valuable insights from vast amounts of information. The Infinite Interact: Multi-Agent Operational Suite is a manifestation of these capabilities, developed as a final project for the CMPE 258 course.

Our project aims to harness the capabilities of Large Language Models (LLMs) and integrate them with LangChain and OpenAI to create a suite of AI agents. These agents are designed to perform a variety of tasks, each addressing specific operational needs. The suite includes:

1. AssemblyAI Agent: This agent is designed to retrieve YouTube videos, download them, generate transcriptions, and allow users to query the transcribed content.
2. PandasAI Agent: This agent facilitates the upload and querying of CSV files, primarily focusing on generating visual plots of the data.
3. Presentation Agent: Tailored for creating PowerPoint presentations from uploaded files, this agent currently supports .ipynb files.

4. README Agent: This agent generates README files from a specified file path, aiding in the documentation process for future projects.
5. Web Scraping Agent: By providing a URL, this agent can scrape web data and allow users to query the retrieved content.

Through these agents, our project aims to demonstrate the versatility and efficiency of AI in automating complex tasks, thereby enhancing the overall workflow and productivity. This report provides a comprehensive overview of the project's objectives, development process, and the specific functionalities of each agent, highlighting their practical applications and potential benefits.

Related Work

The development of AI agents capable of automating various tasks has been an area of significant research and development. Several frameworks and projects have paved the way for the creation of sophisticated AI agents, demonstrating the potential and versatility of AI in different domains.

CrewAI is one of the prominent frameworks in the field of AI agent development. CrewAI focuses on building collaborative AI systems where multiple agents work together to achieve complex goals. This framework emphasizes the integration of various AI technologies to create agents that can communicate, collaborate, and learn from each other. The agents in CrewAI are designed to handle a wide range of tasks, from simple data processing to more complex decision-making processes, making it a versatile tool for developing multi-agent systems.

Other notable AI agent frameworks include:

1. Dialogflow: Developed by Google, Dialogflow is a natural language processing (NLP) platform used for building conversational interfaces such as chatbots, voice

apps, and interactive voice response systems. It allows developers to design and integrate conversational user interfaces into mobile apps, web applications, devices, and bots.

2. OpenAI GPT-3: As one of the most advanced language models, GPT-3 has been instrumental in the development of AI agents capable of understanding and generating human-like text. Its applications range from automated content generation and language translation to coding assistance and more.
3. LangChain: A framework that focuses on building applications with LLMs by chaining together various components. LangChain is particularly useful for creating complex workflows where the output of one task becomes the input for another. It facilitates the development of AI agents that can perform multi-step tasks, integrating different AI capabilities seamlessly.

Our project, Infinite Interact: Multi-Agent Operational Suite, builds upon the principles and advancements demonstrated by these frameworks. By leveraging LangChain and OpenAI, we have developed a suite of AI agents designed to perform a variety of tasks, including video transcription, CSV querying, presentation generation, README file creation, and web scraping. Each of these agents is tailored to address specific operational needs, showcasing the practical applications and benefits of AI in automating diverse tasks.

Supported Data Types

The Infinite Interact: Multi-Agent Operational Suite is designed to handle a diverse range of data types, enabling users to perform various tasks seamlessly. Each AI agent within the suite is equipped to work with specific data formats, ensuring flexibility and utility across different applications. The AssemblyAI Agent retrieves and processes videos from YouTube, downloading them, generating transcriptions, and allowing users to query the transcribed content. The PandasAI Agent supports CSV file uploads, enabling users to query the data within these files and generate visual plots and statistical analyses based on the data. The Presentation Agent currently supports Jupyter Notebook files (.ipynb),

allowing users to upload these files and generate PowerPoint presentations based on their content. The README Agent takes file paths as input and generates README files based on the content found in the specified directory, facilitating the creation of comprehensive project documentation. Lastly, the Web Scraping Agent processes URLs provided by the user, scraping web pages to extract relevant data and allowing users to query this data, making it easy to retrieve information from the web. By supporting various data formats, the Infinite Interact: Multi-Agent Operational Suite demonstrates its versatility and ability to streamline and automate diverse tasks, enhancing productivity and efficiency in different operational contexts.

Methods

AssemblyAI Agent

The AssemblyAI Agent is responsible for handling video transcription tasks. Using LangChain and OpenAI, this agent retrieves a YouTube video by its URL, downloads the video, and processes it to generate a transcription. The user can then interact with the agent to ask questions about the content of the video. This process involves calling the AssemblyAI API to perform the transcription and then utilizing natural language processing (NLP) capabilities of OpenAI's models to handle user queries and provide insightful responses based on the transcribed text.

PandasAI Agent

The PandasAI Agent facilitates data analysis and visualization through CSV files. Users can upload CSV files, which the agent processes using Pandas, a powerful data manipulation library. This agent allows users to perform various queries on the CSV data, generating plots and visualizations to aid in data analysis. The integration with OpenAI enables the agent to understand and execute complex queries, making it a versatile tool for data scientists and analysts.

Presentation Agent

The Presentation Agent streamlines the creation of PowerPoint presentations. Users upload .ipynb (Jupyter Notebook) files, and the agent extracts relevant information to generate slides. This agent leverages the capabilities of LangChain and OpenAI to interpret the notebook content, organize it into a coherent presentation structure, and create visual elements that enhance the presentation. This automation significantly reduces the time and effort required to prepare professional presentations from Jupyter Notebooks.

README Agent

The README Agent is designed to automate the creation of README files for projects. By providing a file path, users can generate comprehensive README files that include project details, setup instructions, and usage guidelines. The agent uses LangChain and OpenAI to parse the project files, extract pertinent information, and format it into a structured README file. This tool ensures consistency and completeness in project documentation, benefiting team members in future assignments and projects.

Web Scraping Agent

The Web Scraping Agent enables users to extract and analyze content from web pages. Given a URL, the agent scrapes the web page and processes the extracted data. Users can then query the data, and the agent, using OpenAI's NLP capabilities, provides answers and insights based on the scraped content. This agent is particularly useful for gathering and analyzing data from various online sources, making it a valuable tool for research and information retrieval.

Data Management

Effective data management is crucial for the successful operation of our AI agents. In the Infinite Interact project, we adopted several strategies to ensure data is handled efficiently and securely.

Data Storage

All project-related data, including source code, transcriptions, CSV files, and generated presentations, are stored in a well-structured directory within a shared repository. Each agent has its own dedicated folder to organize its input and output files, ensuring clarity and ease of access. For instance, the `AssemblyAI Agent` stores transcriptions in a specific folder, while the `PandasAI Agent` maintains uploaded CSV files and generated plots in another.

Conclusion

In conclusion, the Infinite Interact: Multi-Agent Operational Suite exemplifies the transformative potential of AI in automating and enhancing various operational tasks. Leveraging the capabilities of Large Language Models (LLMs) in conjunction with LangChain and OpenAI, our team, Bay Area Rockers, developed a suite of AI agents that cater to diverse user needs, from video transcription and CSV querying to presentation generation, README file creation, and web scraping. Each agent is designed to streamline processes, improve productivity, and provide valuable insights. This project demonstrates how AI can handle different data types and perform complex tasks with ease, showcasing its practical applications in professional and academic settings. The versatility and efficiency of our agents highlight the impact of AI in modern workflows, underscoring the collaborative efforts and innovative spirit of our team. As we continue to explore and expand AI's capabilities, projects like this pave the way for future advancements, enabling more intelligent and efficient tools that can significantly benefit various industries and domains.

References

1. LangChain Documentation. (n.d.). Retrieved from
<https://python.langchain.com/v0.2/docs/introduction/>
2. LangChain API Reference. (n.d.). Retrieved from
[https://api.python.langchain.com/en/latest/langchain_api_reference.html](https://api.python.langchain.com/en/latest/langchain_api_reference.html)
3. OpenAI API Reference. (n.d.). Retrieved from
<https://platform.openai.com/docs/api-reference>
4. Towards Data Science. (2023). A Gentle Intro to Chaining LLMs, Agents, and Utils via LangChain. Retrieved from
<https://towardsdatascience.com/a-gentle-intro-to-chaining-langs-agents-and-utils-via-langchain-16cd385fca81>
5. Reddit. (2023). Chains vs Agents. Retrieved from
[https://www.reddit.com/r/LangChain/comments/14mbyjm/chains_vs_agents/](https://www.reddit.com/r/LangChain/comments/14mbyjm/chains_vs_agents/)
6. Berman, M. (2023). LangChain Overview. [Video]. YouTube. Retrieved from
[https://www.youtube.com/watch?v=9ZhbA0FHZYc&ab_channel=MatthewBerman](https://www.youtube.com/watch?v=9ZhbA0FHZYc&ab_channel=MatthewBerman)
7. HumansdotAI. (2023). An Introduction to AI Agents. Retrieved from
<https://medium.com/humansdotai/an-introduction-to-ai-agents-e8c4afd2ee8f>
8. Towards Data Science. (2023). Create an Agent with OpenAI Function Calling Capabilities. Retrieved from
<https://medium.com/towards-data-science/create-an-agent-with-openai-function-calling-capabilities-ad52122c3d12>

Appendix

The screenshot shows the main dashboard of the Infinite Interact platform. On the left, there's a sidebar with a tree view of project components: main, multimedia, pandasai, presentations, and readme generator. Below this is a "Settings" section with a checkbox for "Enable Text-to-Speech". The main content area features a large heading "Infinite Interact: Multi-Agent Operational Suite" with a small robot icon. A brief description explains the project's goal of developing intelligent agents for document management and extraction. It lists five agents: Multimedia Analysis Agent, DataFrame Assistant Agent, Presentation Creator Agent, Documentation Generator Agent, and WebContent Extractor Agent, each accompanied by a small icon.

The screenshot shows the "Multimedia-AGENT" component of the platform. The sidebar includes "main", "multimedia" (which is selected), "pandasai", "presentations", and "readme generator". The "Settings" section has a "Enable Text-to-Speech" checkbox. Under "File Uploader:", there's a "Choose files" button and a "Drag and drop files here" area with a limit of 200MB per file for MP3, MP4, and MPEG4 formats. A "Browse files" button is also present. The "Number of files uploaded" field shows "0". Below it, a "Pick a color for the answer space" input field contains a small red square. The main content area features a heading "Multimedia-AGENT" with a robot icon. It describes how LLMs and Assembly AI can interact with video and audio files, transcribing them and detecting speakers. It includes a "Youtube Link:" input field and a "Let's start interacting with Multimedia AI!" button. In the bottom right corner, there's a small preview window showing a snippet of the platform's interface.

Let's start interacting with Multimedia AI!

Youtube Link:

Enter a Youtube link
https://www.youtube.com/watch?v=o-z3WvMDP0Q&ab_channel=FOX11LosAngeles

Running youtube_upload(...).



Transcribing video...

Video ready to be communicated with!

Uploaded Document:

Speaker A: Iran's president and foreign minister have both been killed in a helicopter crash. The US State Department now says the deaths are unlikely to have an impact on Iran's actual policy, such as their support for Hamas and their push to provide nuclear weapons. Fox's Rebecca Castor reports from Washington. Speaker B: After searching for hours in the mountains, Iranian soldiers made it to the crash site on Monday, where a helicopter carrying Iranian President Ibrahim Raisi and the country's foreign minister went down. They found no survivors. Speaker D: At this point, there is no evidence of foul play. It looks like an accident. Speaker B: Iranian officials will investigate the crash. And while the country mourns President Raisi's death, analysts say there are many who will not miss him. Speaker C: This Ibrahim Raisi was considered the butcher of Iran, the butcher of Tehran by Iranians for his personal role in one of the greatest mass executions of political prisoners. Speaker B: While Raisi's death is a significant loss, it's not likely to change current policy inside Iran, but it does raise questions about the country's future. Speaker E: The supreme leader, Ayatollah Ali Khamenei, who sets the policy for Iran, and certainly Raisi was a protege and a potential successor to the supreme leader, but he's not determining the policy. Speaker B: Iran is known for backing militant proxy groups like Hamas and Hezbollah. And recently the regime has cracked down on women's rights and there's been a massive surge in government-backed executions. The US State Department says those policies will likely continue. Speaker D: We're going to just trade this terrorist for the next terrorist that the, that the Ayatollahs are going to install. Speaker B: Iran's vice president, Mohammed Makbeh, will serve as interim president. Iranian law requires new elections be held within 50 days. In Washington, Rebecca Castor, Fox News.

main

multimedia

pandasai

presentations

readme generator

webscraper

Settings:

Enable Text-to-Speech

File Uploader:

Choose files

Drag and drop files here
Limit 200MB per file • MP3, MP4, MPEG4

Browse files

Number of files uploaded

0

Pick a color for the answer space

will investigate the crash. And while the country mourns President Raisi's death, analysts say there are many who will not miss him. Speaker C: This Ebrahim Raisi was considered the butcher of Iran, the butcher of Tehran by Iranians for his personal role in one of the greatest mass executions of political prisoners. Speaker B: While Razi's death is a significant loss, it's not likely to change current policy inside Iran, but it does raise questions about the country's future. Speaker E: The supreme leader, Ayatollah Ali Khamenei, who sets the policy for Iran, and certainly Raisi was a protege and a potential successor to the supreme leader, but he's not determining the policy. Speaker B: Iran is known for backing militant proxy groups like Hamas and Hezbollah. And recently the regime has cracked down on women's rights and there's been a massive surge in government backed executions. The US State Department says those policies will likely continue. Speaker D: We're going to just trade this terrorist for the next terrorist that the, that the Ayatollahs are going to install. Speaker B: Iran's vice president, Mohammed Makbeh, will serve as interim president. Iranian law requires new elections be held within 50 days. In Washington, Rebecca Castor, Fox News.

Which president was killed in a helicopter crash?



Recognized:

Which president was killed in a helicopter crash?

AI Response: Iranian President Ibrahim Raisi.

Interaction finished

main

multimedia

pandasai

presentations

readme generator

Settings:

Enable Text-to-Speech

Upload CSV

Drag and drop file here
Limit 200MB per file • CSV

Browse files

PANDAS AI (Chart Generator)

This Agent allows you to use the power of language models to interact with your data. Add a file and start creating visuals and insights using only natural language.

Let's start interacting with PandasAI!

Enter a question

Please upload a CSV file

PANDAS AI (Chart Generator) 🐻

This Agent allows you to use the power of language models to interact with your data. Add a file and start creating visuals and insights using only natural language.

Let's start interacting with PandasAI!

Enter a question

Document Expander (Press button on the right to fold to fold or unfold)

Uploaded Document:

	rank	discipline	phd	service	sex	salary
0	Prof	B	56	49	Male	186,960
1	Prof	A	12	6	Male	93,000
2	Prof	A	23	20	Male	110,515
3	Prof	A	40	31	Male	131,205
4	Prof	B	20	18	Male	104,800
5	Prof	A	20	20	Male	122,400
6	AssocProf	A	20	17	Male	81,285
7	Prof	A	18	18	Male	126,300
8	Prof	A	29	19	Male	94,350
9	Prof	A	51	51	Male	57,800

PANDAS AI (Chart Generator) 🐻

This Agent allows you to use the power of language models to interact with your data. Add a file and start creating visuals and insights using only natural language.

Let's start interacting with PandasAI!

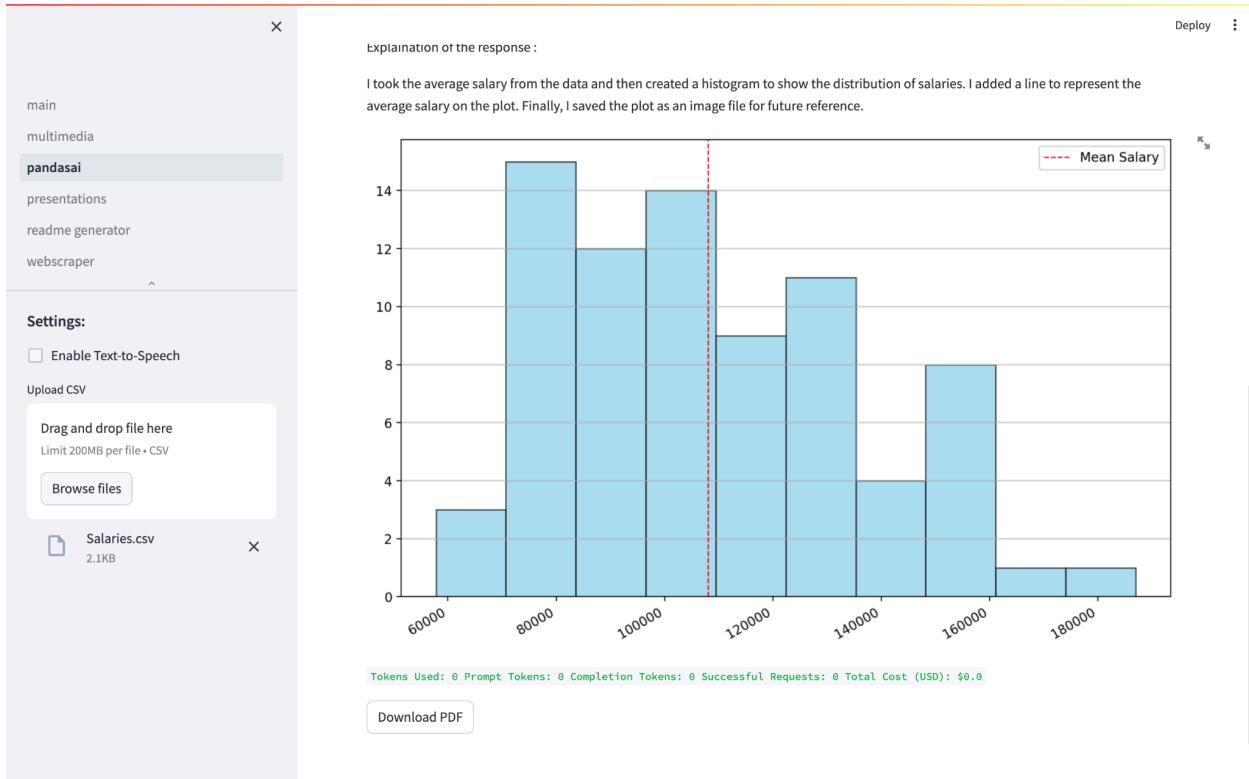
Enter a question

Can you plot the mean distribution of salary?

Document Expander (Press button on the right to fold to fold or unfold)

Uploaded Document:

	rank	discipline	phd	service	sex	salary
0	Prof	B	56	49	Male	186,960
1	Prof	A	12	6	Male	93,000
2	Prof	A	23	20	Male	110,515
3	Prof	A	40	31	Male	131,205
4	Prof	B	20	18	Male	104,800
5	Prof	A	20	20	Male	122,400
6	AssocProf	A	20	17	Male	81,285
7	Prof	A	18	18	Male	126,300
8	Prof	A	29	19	Male	94,350
9	Prof	A	51	51	Male	57,800



PRESENTATION AGENT

Use the power of LLMs with LangChain and OpenAI GPT-4o to scan through your Notebooks and create a Powerpoint presentation. Find information and insight's with lightning speed. Create new content with the support of state of the art language models and and voice command your way through your documents. 🎤

Number of files uploaded
0

Pick a color for the answer space

Deploy ⋮

PRESENTATION AGENT 📈

Use the power of LLMs with LangChain and OpenAI GPT-4o to scan through your Notebooks and create a Powerpoint presentation. Find information and insight's with lightning speed. Create new content with the support of state of the art language models and and voice command your way through your documents. 🎙

Document Expander (Press button on the right to fold or unfold)

Uploaded Document:

A Presentation generated from a Jupyter Notebook

Section 1: Introduction

First we import the needed libraries: `ppsx` and `nbformat`.

```
import nbformat
from ppsx import Presentation
from ppsx.util import Inches
```

Section 2: Function to create a presentation

We define a function that takes a Jupyter Notebook and a PowerPoint template as input and returns a PowerPoint presentation.

```
def read_ipynb_file(file_path):
    with open(file_path, 'r', encoding='utf-8') as file:
        notebook = nbformat.read(file, as_version=4)
```

Deploy ⋮

README GENERATOR 📖

Use the power of LLMs with LangChain and OpenAI to scan through your documents. Find information and insight's with lightning speed. Create new content with the support of state of the art language models and and voice command your way through your documents. 🎙 This LangChain Agent searches a directory for markdown files and generates a README.md file with the most important content.

Let's start interacting with GPT-4o!

Enter the folder path:

Show folder structure

Describe the Project:

main
multimedia
pandasai
readme generator
readme generator

Temperature: 0.10
Recording time: 6

Settings:
 Enable Text-to-Speech

README GENERATOR 📖

Use the power of LLMs with LangChain and OpenAI to scan through your documents. Find information and insight's with lightning speed. 🚀

Create new content with the support of state of the art language models and and voice command your way through your documents. 🎙️ This LangChain Agent searches a directory for markdown files and generates a README.md file with the most important content.

Let's start interacting with GPT-4o!

Enter the folder path:

```
/Users/schumbar/Source/CMPE-258-Final-Project
```

Found 201 files in the directory:

Document Expander (Press button on the right to fold or unfold)

	Filepath	Filename	Fileextension
36	/Users/schumbar/Source/CMPE-258-Final-Project/.git/info	info	
37	/Users/schumbar/Source/CMPE-258-Final-Project/.git/logs	logs	
38	/Users/schumbar/Source/CMPE-258-Final-Project/.git/description	description	
39	/Users/schumbar/Source/CMPE-258-Final-Project/.git/hooks	hooks	
40	/Users/schumbar/Source/CMPE-258-Final-Project/.git/refs	refs	
41	/Users/schumbar/Source/CMPE-258-Final-Project/.git/index	index	
42	/Users/schumbar/Source/CMPE-258-Final-Project/.git/packed-refs	packed-refs	
43	/Users/schumbar/Source/CMPE-258-Final-Project/.git/COMMIT_EDITMSG	COMMIT_EDITMSG	
44	/Users/schumbar/Source/CMPE-258-Final-Project/.git/FETCH_HEAD	FETCH_HEAD	
45	/Users/schumbar/Source/CMPE-258-Final-Project/.git/objects/95	95	

Show folder structure

README GENERATOR 📖

Document Expander (Press button on the right to fold or unfold)

	Filepath	Filename	Fileextension
36	/Users/schumbar/Source/CMPE-258-Final-Project/.git/info	info	
37	/Users/schumbar/Source/CMPE-258-Final-Project/.git/logs	logs	
38	/Users/schumbar/Source/CMPE-258-Final-Project/.git/description	description	
39	/Users/schumbar/Source/CMPE-258-Final-Project/.git/hooks	hooks	
40	/Users/schumbar/Source/CMPE-258-Final-Project/.git/refs	refs	
41	/Users/schumbar/Source/CMPE-258-Final-Project/.git/index	index	
42	/Users/schumbar/Source/CMPE-258-Final-Project/.git/packed-refs	packed-refs	
43	/Users/schumbar/Source/CMPE-258-Final-Project/.git/COMMIT_EDITMSG	COMMIT_EDITMSG	
44	/Users/schumbar/Source/CMPE-258-Final-Project/.git/FETCH_HEAD	FETCH_HEAD	
45	/Users/schumbar/Source/CMPE-258-Final-Project/.git/objects/95	95	

Show folder structure

Describe the Project:

This project has the following functionality:

1. We are able to input a Youtube URL and we get a transcript of the video. We can also query questions regarding the video and an AI Agent will answer those questions.
2. We can upload a CSV file and ask queries regarding that CSV file that will result in plots.

Press ⌘+Enter to apply

The screenshot shows a user interface for a multifaceted AI-powered project. On the left, a sidebar lists repository files: main, multimedia, pandasai, presentations, readme generator (which is selected), and webscraper. Below this are two slider controls: 'Temperature' set to 0.10 and 'Recording time' set to 6. Under 'Settings', there is a checkbox for 'Enable Text-to-Speech'. The main content area has a title 'Project Overview' with a compass icon. It welcomes users to the project and provides a quick rundown of functionalities. A numbered list details the features: 1. YouTube Video Transcription and Q&A, 2. CSV File Analysis, 3. Notebook to Presentation, 4. README Generator, and 5. Web Scraping and Querying. Below this is a section titled 'Prerequisites' with a warning icon, listing requirements: Python 3.8 or higher, pip, and internet connection. The 'Project Setup' section includes steps for cloning the repository and installing dependencies with terminal commands shown in a code block.

Project Overview

Welcome to our multifaceted AI-powered project! This repository offers a suite of tools designed to enhance your productivity and streamline your data processing tasks. Here's a quick rundown of the functionalities:

1. **YouTube Video Transcription and Q&A:** Input a YouTube URL to get a transcript of the video. You can also query questions regarding the video, and an AI Agent will provide answers.
2. **CSV File Analysis:** Upload a CSV file and ask queries about the data. The system will generate relevant plots based on your queries.
3. **Notebook to Presentation:** Upload a `.ipynb` file and automatically create presentations from the notebook content.
4. **README Generator:** Provide a folder structure and generate a comprehensive `README.md` file.
5. **Web Scraping and Querying:** Perform web scraping and query the web-scraped information.

Prerequisites

Before you begin, ensure you have met the following requirements:

- Python 3.8 or higher
- pip (Python package installer)
- Internet connection for web scraping and YouTube transcription

Project Setup

Follow these steps to set up the project on your local machine:

1. **Clone the repository:**

```
git clone https://github.com/yourusername/your-repo-name.git
cd your-repo-name
```

2. **Install dependencies:**