Trial Project: End-to-End Data Extraction, AI Processing, and Automation

Task Description

You are tasked with creating an end-to-end solution that involves web scraping, AI-driven content processing, and automation. Please follow the steps below:

1. Web Scraping

- **Objective:** Write a Python script to extract the **titles** and **summaries** of the latest articles from a news website of your choice (e.g., BBC News, CNN, Reuters).
- **Requirements:**
    - Use **BeautifulSoup** or **Scrapy** for web scraping.
    - Ensure your script efficiently handles HTTP requests and parses HTML content.
    - Include error handling for potential issues like network errors or changes in website structure.

2. AI and NLP Processing

- **Objective:** Paraphrase the extracted summaries to create original content.
- **Requirements:**
    - Utilize an AI model or NLP library such as **OpenAI's GPT-3**, **NLTK**, or **spaCy**.
    - The paraphrased content should maintain the original meaning but use different wording and sentence structures.
    - Ensure the rewritten summaries are coherent and free of grammatical errors.

3. Automation and Integration

- **Objective:** Automate the posting of the paraphrased content to a platform using its API.
- **Requirements:**
    - Choose a platform to post the content, such as:
        - A **Slack channel**
        - A **WordPress site**
        - **GitHub Gist**
        - **Etc.**
    - Write a script to automatically post the paraphrased summaries to the chosen platform.
    - Implement basic error handling to manage network issues or failed posts gracefully.

4. Communication Skills

- **Objective:** Demonstrate your ability to communicate effectively in English.
- **Requirements:**
    - Create a **Loom video** (up to **5 minutes**) where you:
        - Explain your approach and thought process for the project.
        - Demonstrate your script in action.
        - Discuss any challenges faced and how you overcame them.
        - Highlight any decisions or assumptions you made during development.

Deliverables

1. **Python Script(s):**
    - The complete script(s) performing all the tasks outlined above.
    - Ensure your code is well-organized and commented where necessary.
2. **Output File:**
    - A file (e.g., CSV or TXT) containing:
        - The original titles and summaries.
        - The paraphrased summaries.
3. **Loom Video Link:**
    - A URL to your Loom video presentation.
4. **Optional - README File:**
    - A brief document including:
        - An overview of your project.
        - Instructions on how to set up and run your script(s).
        - Any additional notes or considerations.

Instructions

- **Time Allocation:**
    - Please spend no more than **2 hours** completing this project.
- **Programming Language:**
    - Use **Python** for all scripting tasks.
- **Libraries and Tools:**
    - You are free to use any libraries or tools you are comfortable with.
    - Mention all libraries and tools used in your README file or Loom video.
- **API Keys and Access:**
    - If using an API that requires a key (e.g., OpenAI's GPT-3), and you do not have access, you may use open-source alternatives like **NLTK** or **spaCy**.

    - For the posting task, if you don't have access to a platform's API, you can simulate the process or use a mock service.
- **Error Handling:**
    - Include basic error handling to ensure your script can handle exceptions without crashing.

Evaluation Criteria

Your submission will be evaluated based on the following criteria:

Technical Skills

- **Web Scraping:**
    - Ability to extract data accurately and efficiently.
    - Proper use of web scraping libraries.
- **AI and NLP Proficiency:**
    - Effective use of AI/NLP tools to paraphrase content.
    - Quality and originality of the paraphrased summaries.
- **Automation and Integration:**
    - Successful automation of the posting process.
    - Correct implementation of API integrations.

Problem-Solving

- **Error Handling:**
    - Ability to anticipate and manage potential errors.
    - Graceful handling of exceptions and failures.
- **Decision-Making:**
    - Justification of the choices made regarding tools and methods.

Communication

- **Clarity:**
    - Clear and understandable explanations in the Loom video.
    - Effective demonstration of the script's functionality.
- **Language Proficiency:**
    - Proper use of English grammar and vocabulary.
    - Ability to convey technical concepts effectively.

Code Quality

- **Readability:**

    - Clean and well-organized code structure.
    - Use of meaningful variable and function names.
- **Documentation:**
    - Comments explaining key sections of the code.
    - Comprehensive instructions in the README file (if provided).

Submission Guidelines

- **Packaging:**
    - Organize all your files in a single folder or compressed archive (e.g., ZIP file).
    - Name your archive using the following format: `Lastname_Firstname_TrialProject.zip`
- **File Structure:**
    - Ensure all scripts, output files, and the README (if included) are properly named and easy to locate.
- **Loom Video:**
    - Include the link to your Loom video in the README file or in a separate text file named `Loom_Link.txt`.
- **Submission Method:**
    - Reply to this application with your submission attached or provide a download link if the file size is too large.

Additional Notes

- **Focus on Quality:**
    - It's better to submit a high-quality partial solution than a rushed full solution.
- **Resourcefulness:**
    - Feel free to use online resources, but avoid copying code without understanding it.
- **Questions:**
    - If you have any questions or need clarifications, don't hesitate to reach out.

We look forward to reviewing your work. Good luck!