Blackcoffer

Consulting Website: https://blackcoffer.com | https://lsalead.com/

Web App Products: https://netclan.com/ | https://insights.blackcoffer.com/ | https://hirekingdom.com/ | https://workcroft.com/

Mobile App Products: Netclan | Bwstory

Data Extraction and NLP

Test Assignment

Submitted by Shobhit Kumar Singh

Instructions for Running the Source Code

- **Extract Source Code:** Extract the contents of the source_code.zip file.
- **Store Content in Root Directory:** Place all the extracted files and folders into a single directory (root directory). This directory will be the main working directory for the code.
- **Ensure Config.yaml File is Present:** Confirm that the config.yaml and requirements.txt files are present in the root directory.
- **Run main.py:**
  - Open a terminal or command prompt.
  - Navigate to the root directory where the source code is stored.
  - Use pip to install the requirements.txt file
  - Run the main.py file using python

- **Monitor Real-Time Progress:** During execution, a real-time Excel sheet named realtime_sheet.xlsx will be created in the root directory. This file will track the progress of the code.
- **Final Results:** Once the process is complete, the final results will be saved in an Excel file named output.xlsx in the root directory.

Notes:

- During NLTK import or download(), if you are facing any server-related issue then you have to change your proxy setting or manually download NLTK library and its dependencies and save it in their respective directories.
- The code uses relative file paths for configuration and input files. The assumption is that all required files are in the same directory as the source code.
- The code has been developed and tested on a MacBook Pro M1 model. Ensure execution on a similar system to avoid compatibility issues.
- If any path-related or compatibility issues arise, please consider executing the source code on a macOS environment.

Function Comments:

- Each function in the source code is well-commented, providing details like
  - Purpose of the function,
  - Arguments (including its description and its data types),
  - Return type of the function.

Review these comments for a comprehensive understanding of the code's functionality.

Thank you for your time and consideration in reviewing the instructions and source code. In the event of any minor mistakes or oversights, I am open to discussion and ready to address any concerns. Looking forward to your response.

I am passionate about contributing to the world of Machine Learning and AI. If you have further questions or need assistance, feel free to Discuss.
Have a fantastic day!

Best regards,
Shobhit Kumar Singh