

Energy-Moonshot-AI

The Energy Moonshot AI project extracts a KG from UNDP dataset and creates a KG-LLM based pipeline to perform QA tasks. The project contains two sub-modules.

Requirements

The **requirements.txt** files contains all the dependencies to run this project. In the notebooks, pip commands for installation of needed packages are attached but additional packages may be needed depending on the environment.

Additionally, the project requires connection with running **Neo4j** database, and key for **openAI API**. The openAI key used for development is provided by UNDP for the project team and cannot be shared for security reasons.

KG-Extraction Folder

The **KG Extraction Pipeline.ipynb** notebook contains the whole pipeline for KG extraction. Here is a structure of folder:

- **Data:** folder contains the dataset of documents.
- **Entities:** folder contains the files of extracted entities for each document
- **Relations:** folder stores the extracted relations for each document.

KG-LLM Integration Folder

The **Moonshot AI- Backend.ipynb** contains the pipeline for reponse generation. The notebook can run individually wiht the installed dependencies. The files relevant to thesis are:

- **queries.json:** contains the list of queries used for evaluation.
- **LLM Outputs:** contains reponses of the queries by KG-LLM model.
- **KG-LLM Outputs:** contains reponses of the queries by KG-LLM model.
- **excerpts.json:** file temporarily stores the extracted text sections by the semantic model.
- **Autoeval:** folder contains the results of evluation generated by Moonshot AI-Backend.ipynb

Other folders store embeddings and states for semantic search and must not be changed.

Neo4j database:

To run the Neo4j server locally, install the Neo4j Desktop application. Import the given **neo4j.dump file**. Install the APOC Plugin. Run the database and connect with the given connection address by replacing it in the notebooks. Update the user and password as well.

Link for installation of Neo4j Desktop: <https://neo4j.com/download/>

Version of Neo4j: 5.3.0

Plugins required: APOC – v 5.3.0

For further assistance, send an email at:

amnakhkhan.ak@gmail.com