# Installation and User Guide

# Install Dependencies

1. This is designed using python 3.8, using other version may not work
2. CD to the “SystemCode” folder

CD SystemCode

1. Install packages

python -m pip install -r requirements.txt

1. Install Spacy package

python -m spacy download en\_core\_web\_sm

# Running Backend (Article & Feature Extraction)

Backend Script is called *MainBackEndProcessing.py*. Run this script to start. Offline mode is available as the tool relies on google news to query news articles. Users will get blocked if too many queries are performed within a short period of time.

* Run test mode (test dataset used to develop the tool)

*Note: Test mode will not write data to db*

python MainBackEndProcessing.py -M test

* Run production mode (live, using today's article). Use -D to run on specific day of articles (*Note: running in production mode requires internet connection*)

python MainBackEndProcessing.py -M prod

# TO run on specific day

python MainBackEndProcessing.py -M prod -D 2023-04-11

* Run demo mode (using 11 April 2023 dataset)
* Online Mode (*Note: running in production mode requires internet connection*)

python MainBackEndProcessing.py -M prod

* Offline Mode

python MainBackEndProcessing.py -M prod -O

The final output is input into database under folder "KnowledgeBase\ApolloDM.db" and its snapshot is located at "\output"

# Starting the Web Application

Run *app.py* script

python app.py

Web is served on <http://127.0.0.1:8050/>

# Repopulating Knowledge Base and Retraining Model

In case the existing Neo4J AuraDB is not working or an update of the knowledge base is required, there are scripts prepared under the “TrainModelOrBuildKB” folder to repopulate the knowledgebase and machine learning models.

The table below provides a description of the scripts that were prepared:

|  |  |
| --- | --- |
| Script Name | Description |
| CountryTravelersNetworkBuilding.ipynb | Script to repopulate travellers network graph to neo4j |
| Disease NER Custom Models.ipynb | Script to retrain disease NER model |
| DiseaseTopicClassification-SVMContextualizedEmbeddings.ipynb | Script to retrain disease topic classification model |
| Diseases Knowledge Base Building.ipynb | Script to repopulate disease knowledge base |
| RelevanceScoreModel.ipynb | Script to retrain relevance score model |