Instructions on How to Run Term Project

Google Colab:
1. If using Google Colab, download the dataset from Kaggel https://www.kaggle.com/datasets/lakshmi25npathi/imdb-dataset-of-50k-movie-reviews?select=IMDB+Dataset.csv to your local drive and set the data file path in the pyscript to your local drive where the file is stored.
2. Run the pyscript with one argument (name of output file)
   run_Term_Project_Sentiment_Analysis.py   <name of output file>

Google Cloud:
Use google cloud spark cluster and submit a job with the following:

Google Cloud link to pyspark script:
gs://muniba-met-cs-777/CS777-Term-Project-Sentiment-Analysis/run_Term_Project_Sentiment_Analysis.py

Google Cloud link to store output:
gs:// muniba-met-cs-777/CS777-Term-Project-Sentiment-Analysis/output/

Google Cloud link to dataset:
gs://muniba-met-cs-777/CS777-Term-Project-Sentiment-Analysis/IMDB_Dataset.csv

Data Set Link:
https://www.kaggle.com/datasets/lakshmi25npathi/imdb-dataset-of-50k-movie-reviews?select=IMDB+Dataset.csv