Backend:

1. RetrieveData.py: Acting as the API to read in and retrieve the movie info from the dataset by providing the path to the dataset
2. DataPreProcessing: Acting as the API to preprocess the read-in data before performing queries
   1. It’s been separated from 1 to provide a loose coupling and to provide the original data presenting to user when query is ended
3. Stopwords removal will only be applied to the spot attribute as other attributes are vital
4. Stemming and stop words removal only applied to plot as other attributes such as title, director, release dates, keywords, certificates are too important to undergo such process.