

Billboard Lyrics Finder

CS F469 - Information Retrieval | Vector Space Model | Domain Specific Search Engine

Language Used: Python

Working Model:

- The search engine can be started by running **app.py**. (python app.py)
- The entire corpus is preprocessed and the tf-idf Table is produced.
- Enter the words / lyrics of the song you want to search for.
- After each search, the top 10 documents that matches with the query are displayed.
- The results will be random if the searched words are not present in any document in the corpus.
- On clicking the Document, one can view the entire document with details like 'artists' etc.

Requirements/Installation:

To run the following code, Flask and nltk have to be readily installed.

- Flask can be installed by following the documentation in the below link.
[Flask_installation](#)
- nltk can be installed using 'nltk.download()' in a python shell or by including it in the code.
- **Libraries:** flask, nltk, csv, wtforms, math

To run the code in Python environment:

- Change directory to the folder having the data set and type the command
“ cd /path/to/project/folder/ “
“ python3 main.py ”
- Now click on the link provided (local server) to access the GUI of the search engine. Then, type your query and find the top 10 relevant results. By clicking on the document name, you can get the details of the document.