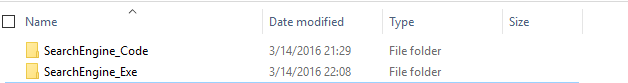
# **Search Engine Program Execution**

The user can choose to either run the program directly or explore the codebase and execute the program from IDE. IDE provide more features to user in terms of error messages or exceptions.

Initial Setting: User should navigate to the folder where the user downloaded the zip file. Unzip it and it should look similar to following image. The folder SearchEngine\_Code contains the .java files or the program code and the SearchEngine\_Exe folder contains the executable jar file.



## **Running the Program**

To run the program the user need to go inside the Executable directory and then select the jar file **SearchEngine.jar**. The user simply needs to double click on the jar file to run the program. User need to make sure that JRE 1.7 is installed on the system.

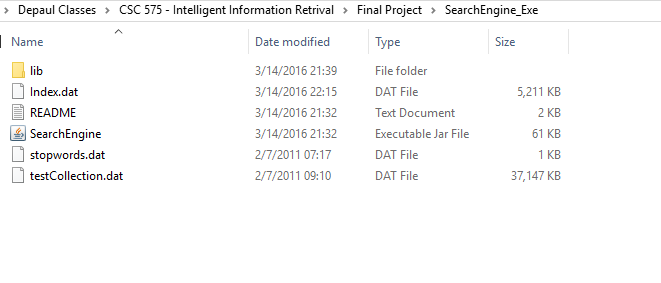
The Program will start executing. The program will display status that the crawler and the indexer are running and the user needs to wait for them to complete processing.

Once the program has finished processing the user could enter the query term and the search box and get the top documents that are relevant to the query. The relevancy matrix used here is Cosine Similarity.

The progress about the system is also visible in the output area located at the bottom of program window.

Please make sure that the two files “stopwords.dat” and “testCollection.dat” are available in the folder.

The directory structure should look similar to following:



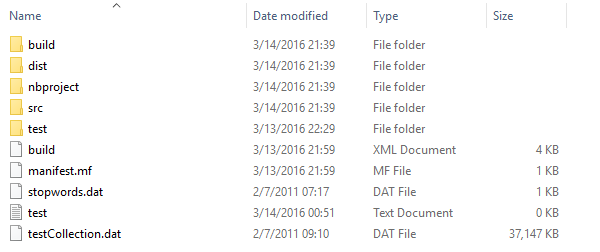
## **Exploring the Code & Run Program Using IDE**

The code was built using the Netbeans IDE version 8.1. I would recommend user to use the same IDE to explore the code. The user could also use other IDE of their choice.

To Explore the code user need to import the code into Netbeans by opening the project in Netbeans. To do this user need to select the File -> Open Project and then navigate to the folder where they have downloaded the project.

The user could easily explore the code and run the code from here.

The structure of directory should be



## Sample Program Execution

When user will first launch the program the program will create a index by running the crawler over the document corpus. The user should see the following screen during this process. User need to wait until this process completes.

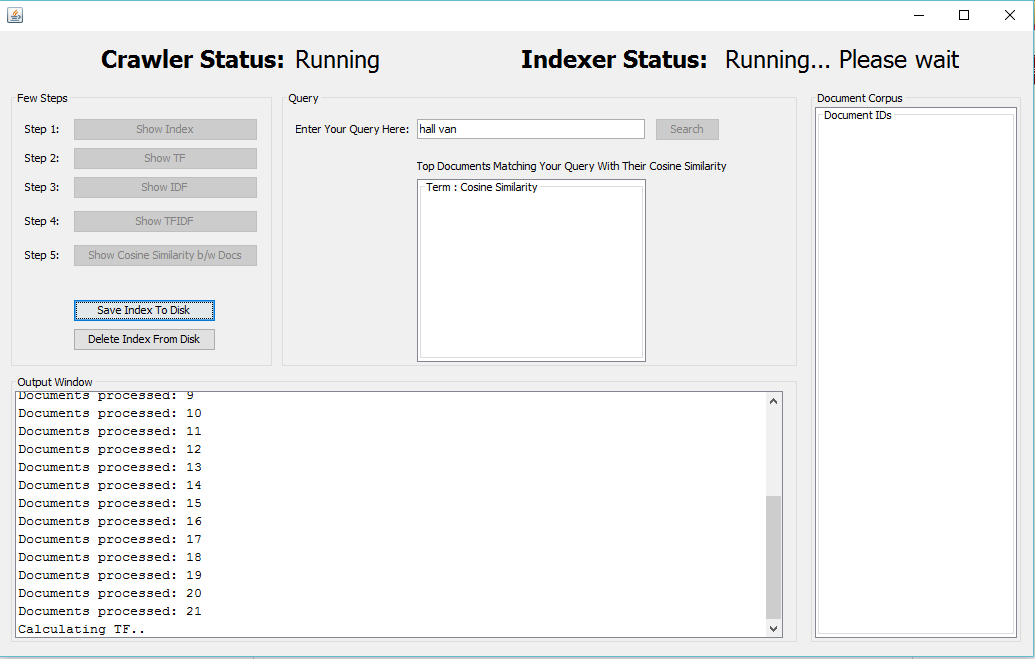
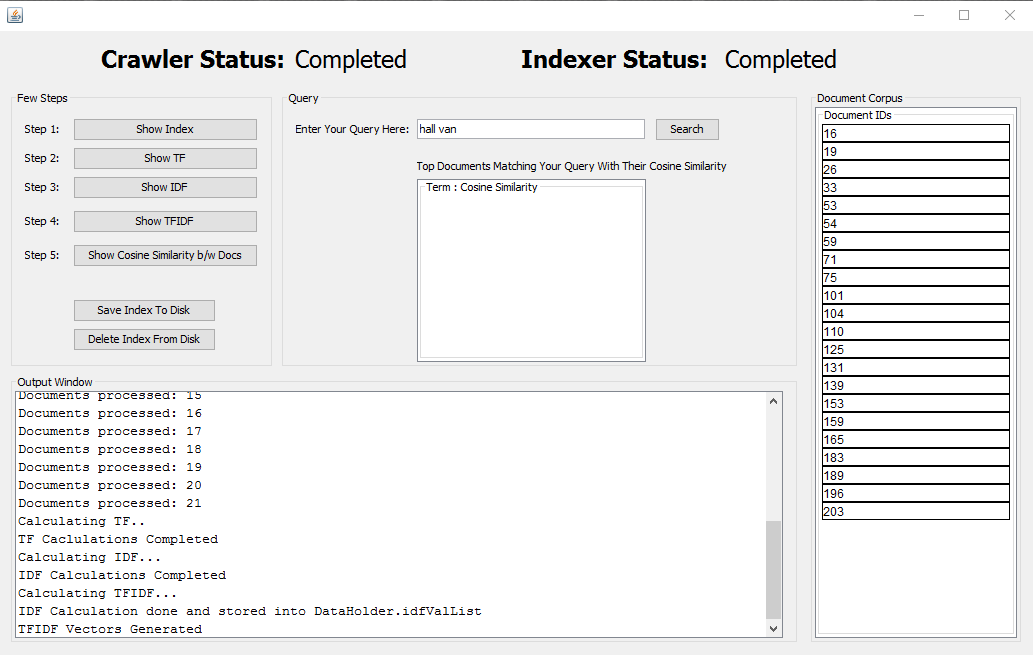


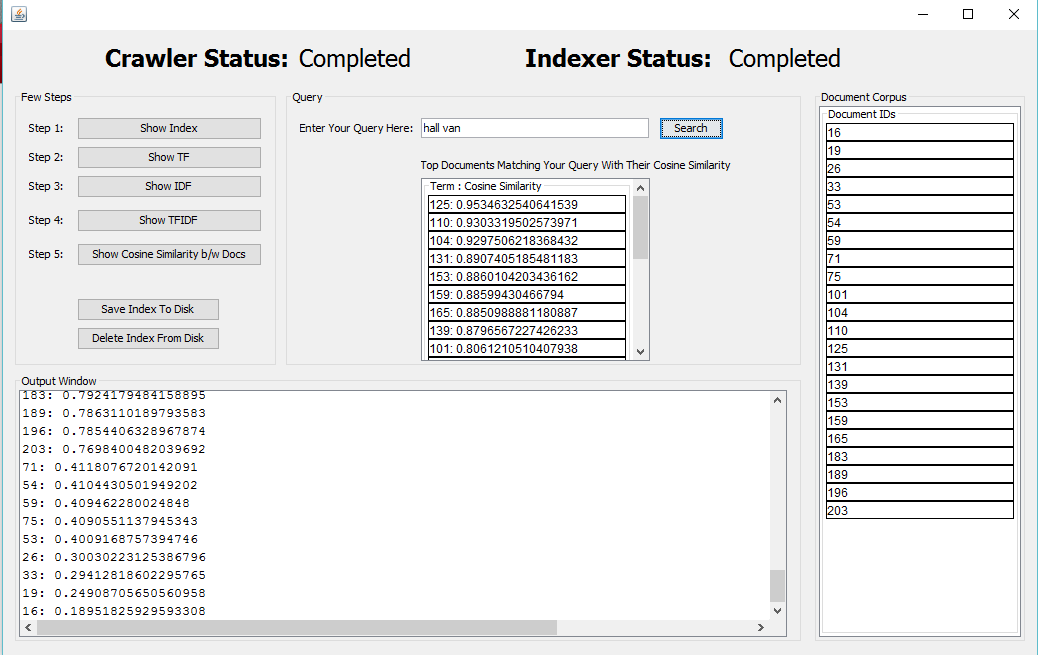
Figure : Screen when program is stared

Once the processes gets completed user will notice that the search button becomes active and the output window has some messages. The Indexer status will be shown as completed.



Querying/ Searching

The user need to type the query in the textbox and click on search button to search for documents that are related to the query. The output will be displayed to user as a list of docs with decreasing relevance. A sample query is already typed in box for user. The result of the query is below.



Other Options for user:

Show Index button : The user could also explore the generated index by clicking on Show Index button. This will show the index that was constructed by the indexer.

Show TF button: This button will show the user TF of a selected document.

Show IDF: This option will show the user IDF values in the output window.

Show TFIDF button: Clicking on this button will show the user the TFIDF values.

Show Cosine Similarity b/s Docs button: This button will show user the cosine similarities between the document corpus.

Save Index to Disk button: Clicking on this button will save index to disk for reloading into program. The user will save time and computing power by reloading the prebuilt index. When the user starts the program the index does not needs to be constructed again.

Delete Index from Disk button: This option will delete the index file from the disk. When user will restart the program will again need to rebuild index.