**NAME: Saloni Kalsekar**

**FINAL PROJECT**

**For completing the project :**

We use all the mentioned requirements in the project.

We will be importing all the webpages mentioned. And extracting the data in the “p” that is the paragraph tag. And then taking each word and searching it in the web pages and also implementing ranking.

First we take three webpages as an input.

A **Trie data structure** is created in the Trie and Trie Structure java files

The Trie operations are performed.

Using the Trie we store the data and search the required word specified in the input.

We use one more data structure hash map for the Ranking

The **inverted index** is the main data structure of our search engine.

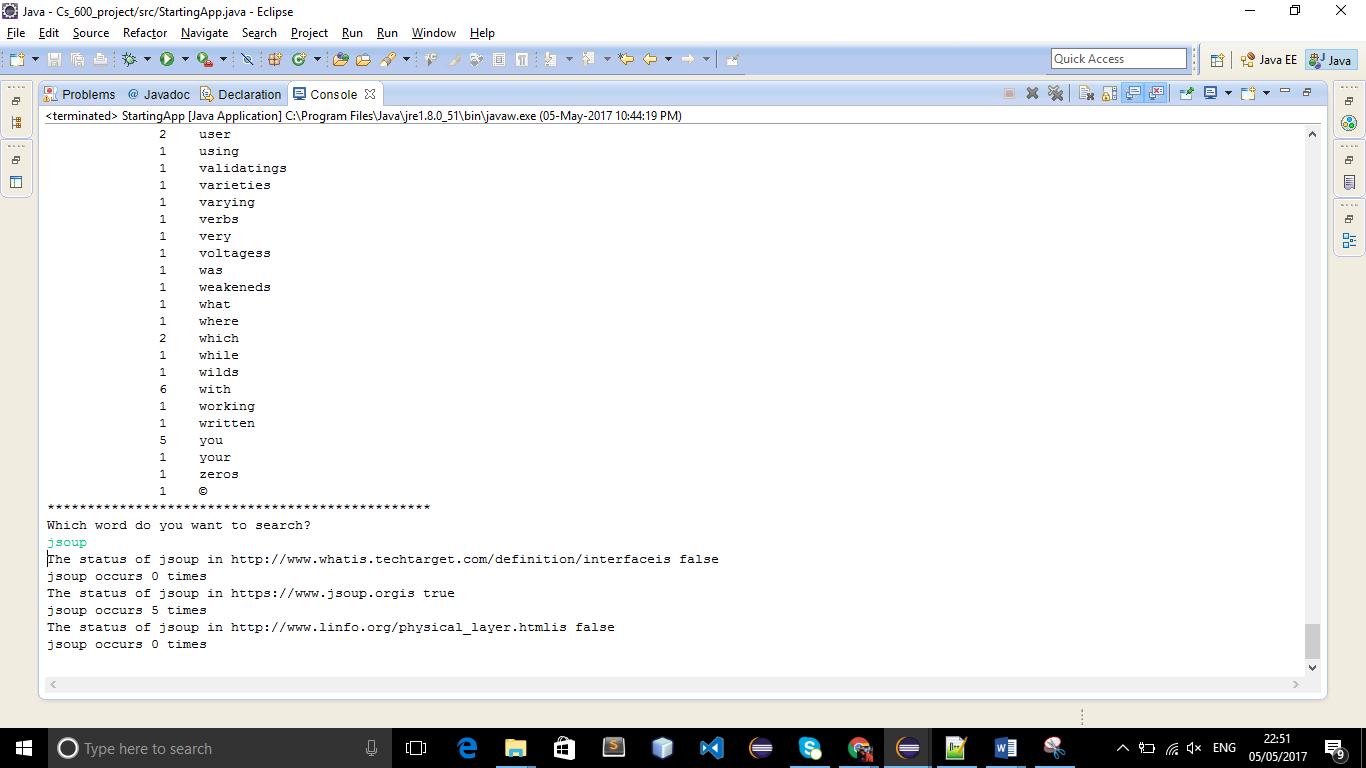
The **map** has a key value pair where the key is the word and the count is the value. We take the count and display the occurences of all the words in the web pages provided.

The **page rank** is high if the occurrence of a word is high in a page. We also display the occurrences in each link.

Below is the output of one of the word searched along with the occurences of few words.

The stop words are removed by comparing the words from the **stopwords**.txt file .

The main output file is displayed in the text file.



For implementing first we need to include the jsoup jar file

And then start by running the StartingApp.java file