Information Retrieval 8130

Assignment 7

Submitted by: Hasan Mashrique

**Due Date:** 11/18/2022

**Problem 1:**

The code does the following tasks:

1. Reads the preprocessed files (edit file location from the code before running) from HomeWork 6 and creates the inverted index (Files link given below)
2. Gets the user query from code (put query in **query** variable in code)
3. Gets all the files that contains the query words
4. Compares each file with query and computes cosine similarity value
5. Stores the file and its cosine similarity value in a dictionary (**sorted\_dic**)
6. Sorts the **sorted\_dic** in descending order and shows the file(url) with similarity value

Run the file to see the output.

The link to all the saved vocabulary files is here:

<https://www.dropbox.com/s/dbr6h74wbm450z9/ProcessedFiles.zip?dl=0>

**Problem 2:**

(Solution with UI will be shown later)