SCHOLARSHIP SEARCH ENGINE

Group Members

Shreyansh Dalwadi (110097536)
Karan Vishavjit (110099867)
Manjinder Singh (110097177)
Neel Pandya (110095825)
Harbhajan Singh (110100089)

Date

March 26, 2023

Course title

Advanced Computing Concepts - COMP8547-1-R-2023W

Professor Name

Dr. Mahdi Firoozjaei

Graduate Assistant

Roisul Islam Rumi

University of Windsor

University of Windsor

CONTRIBUTOR NAME –
SHREYANSH DALWADI

PART CONTRIBUTED –
SORTING, DELETION

Overview of Project

- Scholarship opportunities are a great way to finance one's education. However, finding the right scholarship can be a daunting task. With the vast amount of information available online, it can be overwhelming to sift through all the data and find the most relevant scholarships.
- The scholarship web search engine uses jsoup to crawl web pages and download HTML files. It then reads these files, filters out stop words, and converts them to text files. Once the text files are created, the user inputs a keyword, which is autocorrected using edit distance. The engine then searches for patterns in the downloaded text files using the Boyer-Moore algorithm to find occurrences of the keyword.

Contribution Made by Shreyansh Dalwadi
(110097536)

Sorting Functionality

- SortFile Java class provides a method named **sort()** that is use for sorting a hashtable of HTML links based on their ranking value. The ranking value shows how relevant the HTML link is to a particular search query.
- The method has two parameters:
- "htmlLinks": a hashtable that contains the HTML links as string
- "count": a hashtable that contain the count as integer
- For Sorting , it firstly convert htmlLinks , hashtable to an ArrayList of Map.Entry objects. This is done by using the "entrySet" method on the hashtable, which returns a set of entries.. The "ArrayList" method is used to create a new ArrayList of Map.Entry objects from the set of entries.
- Once the **ArrayList** has been created, the "Collections.sort" method is called to sort the ArrayList in descending order based on the ranking value. Comparator object is used to the sort method, which compares the objects of the entries.
- Lastly, the method prints out the top 10 web pages containing results related to the search query. This is done by iterating through the sorted ArrayList and printing out the key (HTML link) and value (ranking) of the first 10 entries.
- Overall, this class provides a way to sort a hashtable of HTML links based on their ranking values and display the top results in a readable format.

Delete Functionality

- HelperMethods.Java , file is used for deleting all the txt and Html files in two directories,.It contains a method **deleteFiles().**

- The **deleteFiles()** method firstly, creates a File object for the "txt" and "html" directory, which contains the text files to be deleted. then it checks whether, the directory exists using the "exists" method. If the directory exists, deletes the file through iterating the array .

Output

Fig 1 : Unsorted Data

Fig 2: Sorted Data

References

- Lab files provided by Professor Mahdi Firoozjaei
- Overview of Project referred from Group PPT of Scholarship Search Engine.