File Retrieval Engine, Application Layer

Performance Evaluation

Dataset(s)	Indexing Time	Indexing Throughput
Dataset1	22.070 seconds	5.80 MB/s
Dataset2	137.314 seconds	3.73 MB/s
Dataset3	389.859 seconds	5.25 MB/s

OUTPUT:

Dataset1:

Search for: Indian and mathematics. It took 0.97 seconds to search through dataset.

```
[INFO] Scanning for projects...
[INFO]
[INFO] --------------------< csc435.app:FileRetrievalEngine >-------------------
[INFO] Building FileRetrievalEngine 1.0-SNAPSHOT
[INFO] --------------------------------[ jar ]----------------------------------
[INFO]
[INFO] --- exec-maven-plugin:3.1.0:java (default-cli) @ FileRetrievalEngine ---
*****  Choose from the below options to use the File Retrieval Engine application  *****
-------------------------------------------------------------------------------------

  Option 1 : Enter Folder path for search : index <FolderPath>

  Option 2 : Enter word to search in specified file : search <word_to_search>

  Option 3 : Exit the application : quit
-------------------------------------------------------------------------------------

Enter your Option: index /mnt/shared/dataset1
Completed indexing 134200314 bytes of data with 4 worker threads
Completed indexing in 22.070 seconds
Indexing throughput: 5.80 MB/s
-------------------------------------------------------------------------------------

  Option 1 : Enter Folder path for search : index <FolderPath>

  Option 2 : Enter word to search in specified file : search <word_to_search>

  Option 3 : Exit the application : quit
-------------------------------------------------------------------------------------

Enter your Option: search indian and mathematics
Search results (top 10 out of 182):
* /mnt/shared/dataset1/folder7/folderA/Document10956.txt:168
* /mnt/shared/dataset1/folder2/folderA/Document10366.txt:129
* /mnt/shared/dataset1/folder6/Document10891.txt:114
* /mnt/shared/dataset1/folder1/Document10046.txt:98
* /mnt/shared/dataset1/folder7/Document10929.txt:84
* /mnt/shared/dataset1/folder4/Document10669.txt:77
* /mnt/shared/dataset1/folder2/Document1025.txt:76
* /mnt/shared/dataset1/folder6/Document10765.txt:76
* /mnt/shared/dataset1/folder8/Document11151.txt:60
* /mnt/shared/dataset1/folder7/folderD/Document11034.txt:56
Search completed in 0.97 seconds
-------------------------------------------------------------------------------------
```

Dataset2:

Search for: Sterling. It took 88.43 seconds for searching.

```
[INFO] Scanning for projects...
[INFO]
[INFO] ---------------< com.fileretrieval:FileRetrievalEngine >----------------
[INFO] Building FileRetrievalEngine 1.0-SNAPSHOT
[INFO] --------------------------------[ jar ]---------------------------------
[INFO]
[INFO] --- exec-maven-plugin:3.1.0:java (default-cli) @ FileRetrievalEngine ---
No text files found in the folder.
*****  Choose from the below options to use the File Retrieval Engine application  *****
----------------------------------------------------------------------------------------

 Option 1 : Enter Folder path for search : index <FolderPath>

 Option 2 : Enter word to search in specified file : search <word_to_search>

 Option 3 : Exit the application : quit
----------------------------------------------------------------------------------------

[Enter your Option: index /mnt/shared/dataset2
Completed indexing 536991226 bytes of data with 1 worker threads
Completed indexing in 137.314 seconds
Indexing throughput: 3.73 MB/s
----------------------------------------------------------------------------------------

 Option 1 : Enter Folder path for search : index <FolderPath>

 Option 2 : Enter word to search in specified file : search <word_to_search>

 Option 3 : Exit the application : quit
----------------------------------------------------------------------------------------

[Enter your Option: Search Sterling
Search results (top 10 out of 171):
* /mnt/shared/dataset2/folder3/Document11259.txt:196
* /mnt/shared/dataset2/folder2/Document1085.txt:178
* /mnt/shared/dataset2/folder6/Document13583.txt:29
* /mnt/shared/dataset2/folder1/Document1025.txt:12
* /mnt/shared/dataset2/folder6/Document13660.txt:12
* /mnt/shared/dataset2/folder6/Document13559.txt:10
* /mnt/shared/dataset2/folder4/Document1224.txt:8
* /mnt/shared/dataset2/folder1/Document10386.txt:7
* /mnt/shared/dataset2/folder7/folderD/Document14831.txt:7
* /mnt/shared/dataset2/folder3/folderA/Document11504.txt:5
Search completed in 88.43 seconds
----------------------------------------------------------------------------------------
```

Dataset3:

Search for: Evolution. It took 218.25 seconds for searching.

```
Picked up JAVA_TOOL_OPTIONS: -Xms2g -Xmx8g
[INFO] Scanning for projects...
[INFO]
[INFO] --------------------< csc435.app:FileRetrievalEngine >--------------------
[INFO] Building FileRetrievalEngine 1.0-SNAPSHOT
[INFO] --------------------------------[ jar ]---------------------------------
[INFO]
[INFO] --- exec-maven-plugin:3.1.0:java (default-cli) @ FileRetrievalEngine ---
*****  Choose from the below options to use the File Retrieval Engine application  *****
-------------------------------------------------------------------------------------

 Option 1 : Enter Folder path for search : index <FolderPath>

 Option 2 : Enter word to search in specified file : search <word_to_search>

 Option 3 : Exit the application : quit
-------------------------------------------------------------------------------------

Enter your Option: index /mnt/shared/dataset3
Invalid folder path: /mnt/shared/dataset3
-------------------------------------------------------------------------------------

 Option 1 : Enter Folder path for search : index <FolderPath>

 Option 2 : Enter word to search in specified file : search <word_to_search>

 Option 3 : Exit the application : quit
-------------------------------------------------------------------------------------

Enter your Option: index /mnt/shared/dataset3
Completed indexing 2144394816 bytes of data with 4 worker threads
Completed indexing in 389.859 seconds
Indexing throughput: 5.25 MB/s
-------------------------------------------------------------------------------------

 Option 1 : Enter Folder path for search : index <FolderPath>

 Option 2 : Enter word to search in specified file : search <word_to_search>

 Option 3 : Exit the application : quit
-------------------------------------------------------------------------------------

Enter your Option: search Evolution
Search results (top 10 out of 610):
* /mnt/shared/dataset3/folder4/Document1909.txt:276
* /mnt/shared/dataset3/folder1/Document1043.txt:143
* /mnt/shared/dataset3/folder2/folderA/Document14586.txt:106
* /mnt/shared/dataset3/folder7/folderD/Document30134.txt:88
* /mnt/shared/dataset3/folder4/Document19321.txt:74
* /mnt/shared/dataset3/folder2/Document13407.txt:57
* /mnt/shared/dataset3/folder3/folderA/Document17194.txt:53
* /mnt/shared/dataset3/folder2/Document1561.txt:51
* /mnt/shared/dataset3/folder4/Document200.txt:48
* /mnt/shared/dataset3/folder7/folderD/Document30610.txt:42
Search completed in 218.25 seconds
-------------------------------------------------------------------------------------
```