**Assignment 3 – Page Rank**

**Name:** Sharath Chandra Bagur Suryanarayanaprasad

**Email:** [sbagursu@uncc.edu](mailto:sbagursu@uncc.edu)

Student ID: 800974802

The assignment consists of 4 java files FetchN.java, InitPageRank.java, GeneratePageRank.java and Sort.java

FetchN.java – computes the total number of documents

InitPageRank.java- computes and assigns initial Page Rank to all documents

GeneratePageRank.java – Computes page rank by iterating 10 times by distributing the current page rank to all the outlinks.

Sort.java – Sorts the results based on scores of all the documents.

**Execution Instructions:**

Place all the above java files in a folder. Create a folder named input and place all the input files in this folder.

Open terminal and run the below commands from the above folder location-

hadoop fs -rm -r input

hadoop fs -put input/\* input

Run the below commands to get the output from hdfs

rm -r outputs

mkdir -p outputs

Run the below commands to build and execute-

rm -r build

mkdir -p build

rm pagerank.jar

javac -cp /usr/lib/hadoop/\*:/usr/lib/hadoop-mapreduce/\* FetchN.java InitPageRank.java GeneratePageRank.java Sort.java -d build -Xlint

jar -cvf PageRankWiki.jar -C build/ .

hadoop jar pagerank.jar org.myorg.Sort input output

hadoop fs -get output/part-r-00000 outputs/PageRankOutput.out

**Requirements -** The program does require a minimum of Java 8. It is tested in the environment of java version 1.8.0\_91