Readme

Hadoop Commands:

hdfs dfs -put ~/Data/ ~/input

hadoop com.sun.tools.javac.Main Pairs.java

jar cf Pairs.jar Pairs\*.class

hadoop jar Pairs.jar Pairs ~/input/Data ~/op\_pairs

hdfs dfs -get ~/op\_pairs

The data file contains the tweets data, which are analyzed and the output is stored in op\_pairs folder.

The op\_pairs folder has the output and the Data folder has the input.

The output is in the form of word= ‘word’ neighbor= ‘neighbour’ frequency

{word=["1","de] neighbour=[capoli]} 1

{word=["1","de] neighbour=[seminude]} 1