CS 6350

ASSIGNMENT 3

Names of students in your group: 2

Sriprasath Gujuluva Parthasarathy(sxg180154)

Shariq Ali(sxa190016)

Number of free late days used: 2  
Note: You are allowed a **total** of 4 free late days for the **entire semester**. You can use at most 2 for each assignment. After that, there will be a penalty of 10% for each late day.

Please list clearly all the sources/references that you have used in this assignment.

1. Data Intensive Text Processing using MapReduce
2. <https://developer.twitter.com/en/portal/projects-and-apps>
3. <https://kafka.apache.org/quickstart>
4. <https://docs.microsoft.com/en-us/windows/wsl/install-win10>
5. https://spark.apache.org/docs/2.2.0/ mllib-evaluation-metrics.html
6. <https://github.com/databricks/spark-deep-learning>
7. <https://www.elastic.co/downloads/>
8. <https://www.scala-sbt.org/>
9. <http://bahir.apache.org/docs/spark/current/spark-streaming-twitter/>
10. <https://github.com/harpribot/corenlp-scala-examples>
11. <https://databricks-prod-cloudfront.cloud.databricks.com/public/4027ec902e239c93eaaa8714f173bcfc/1233855/3233914658600709/588180/latest.html>
12. <https://github.com/shekhargulati/52-technologies-in-2016/tree/master/03-stanford-corenlp>
13. <https://github.com/nielsutrecht/spark-sentiment>
14. <https://snap.stanford.edu/data/index.html#socnets>

Part 1

Topic: topicA

Index: mytopicindex

Part 2

Dataset: web-Google.txt