

# CS5543 Real-Time Big Data Analytics

## LAB ASSIGNMENT #2

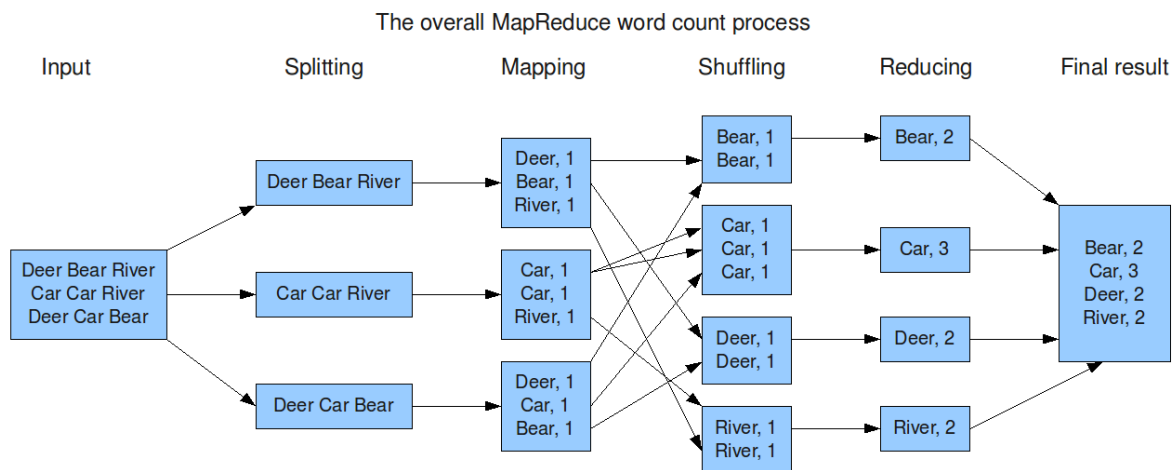
Lab Submission through the following link before 9/7 (W) 11:59PM.

<https://goo.gl/forms/HGJ0VR8U2dH1K9Yv1>

### 1) Spark Programming:

Write a spark program with an interesting use case using text data as the input and program should have at least Two Spark Transformations and Two Spark Actions.

Present your use case in map reduce paradigm as shown below (for word count).



Submit your Github project for Lab Assignments with source code, input and output data and wiki page containing the report for your lab Assignment including map reduce diagram.