

CS5542 Big Data Apps and Analytics

LAB ASSIGNMENT #5

Lab Submission before 2/24 (W) 11:59PM.

Submit a short lab report including screenshots and source code & data to GitHub. Post your GitHub URL through Lab 5 form at <http://goo.gl/forms/tZr32L8P3h>

1. Spark and Smartphone/Watch Application
Implement a smart application with big data analytics related to your project showing the collaboration between Spark and Smart Apps. Note: For the big data analytics part, you can use either SparkR or Spark MLlib.
2. SparkR – Extra Credit
Install Spark and Configure SparkR. Execute Word Count and K-Means Clustering algorithms using your own data (relevant to your project) with SparkR. Write a report including your algorithms and result screenshots. Modify the K-Means SparkR implementation for K-Medoids algorithm. Reference: <https://amplab-extras.github.io/SparkR-pkg/>