**CSEE5590-0001/490-0003: Big Data Programming**

**Lesson Plan # 2**

**ICP Feedback and Submission Link :** [**https://docs.google.com/forms/d/e/1FAIpQLScaeLhVS92maF9Z9exB-5dw2PdPAiljbAeb92zb3jAe9hqVHg/viewform?usp=sf\_link**](https://docs.google.com/forms/d/e/1FAIpQLScaeLhVS92maF9Z9exB-5dw2PdPAiljbAeb92zb3jAe9hqVHg/viewform?usp=sf_link)

**For Online students:** [**https://docs.google.com/forms/d/e/1FAIpQLScaeLhVS92maF9Z9exB-5dw2PdPAiljbAeb92zb3jAe9hqVHg/viewform?usp=sf\_link**](https://docs.google.com/forms/d/e/1FAIpQLScaeLhVS92maF9Z9exB-5dw2PdPAiljbAeb92zb3jAe9hqVHg/viewform?usp=sf_link)

**Lesson Title: *Apache Spark II***

**Lesson Description: *Apache Spark Practice***

**Lesson Overview:**

Apache Spark is a unified analytics engine for big data processing, with built-in modules for streaming, SQL, machine learning and graph processing.

**In class exercise:**

1. ***Merge Sort Algorithm***

Merge Sort is a Divide and Conquer algorithm. It divides input array in two halves, calls itself for the two halves and then merges the two sorted halves. The merge () function is used for merging two halves. The merge (arr, l, m, r) is key process that assumes that arr[l..m] and arr[m+1..r] are sorted and merges the two sorted sub-arrays into one.



Create a Map-Reduce Program to perform Merge-Sort Algorithm in Spark.

1. ***DepthFirst Search***

Implement Depth First Search in Graph in Apache Spark

<https://medium.com/basecs/deep-dive-through-a-graph-dfs-traversal-8177df5d0f13>

**ICP Guidelines (In Class Students):**

1. ICP Submission is in pairs of two students.
2. Once completed, must be presented to TA or Instructor before the completion of the class
3. Submission after class is considered as late submission. (Check the late submission policy in the syllabus)

ICP Code with brief explanation in wiki should be pushed to GitHub (submit your screenshots as well. The screenshot should have both the code and the output)Submit GitHub Link through the Feedback Form (<https://docs.google.com/forms/d/e/1FAIpQLScaeLhVS92maF9Z9exB-5dw2PdPAiljbAeb92zb3jAe9hqVHg/viewform?usp=sf_link>**)**

**Submission Guidelines (for online students):**

1. Submit your source code and documentation to GitHub and represent the work through wiki page properly with detailed explanation (submit your screenshots as well. The screenshot should have both the code and the output)
2. Comment your code appropriately.
3. Submit a brief demo video 2-3 min showing your assignment with a voice over explaining your work through the Submission Link.
4. Use the following Google link to submit your assignment

***Cheating, plagiarism, disruptive behavior and other forms of unacceptable conduct are subject to strong sanctions in accordance with university policy. See detailed description of university policy at the following URL:*** [*https://catalog.umkc.edu/special-notices/academic-honesty/*](https://catalog.umkc.edu/special-notices/academic-honesty/)