

CSEE5590-0001/490-0003: Big Data Programming

Spark: Lesson Plan # 6

ICP Feedback and Submission Link: <https://forms.gle/xMAmr3zATrtMG5cX7>

Lesson Title: *Graph Frames and GraphX Algorithms*

Lesson Description: *Distributed Collection of Data*

Lesson Overview:

- Graph frames
- Algorithms
- Implementations
- References

In Class Exercise

Dataset:

<https://umkc.box.com/s/vaji06fsdw8bmajg3ndh62v2o1hqpvv>

Graph Frames in Pyspark / Scala

Consider the datasets attached above:

References:

<https://stackoverflow.com/questions/37146816/spark-scala-graphx-shortest-path-between-two-vertices>

<http://note.yuhc.me/2015/03/graphx-triangle-count-label-propagation/>

Part – 1:

1. Import the dataset as a csv file and create data frames directly on import than create graph out of the data frame created.
2. Triangle Count
3. Find Shortest Paths w.r.t. Landmarks
4. Apply Page Rank algorithm on the dataset.
5. Save graphs generated to a file.

Bonus:

1. Apply Label Propagation Algorithm
2. Apply BFS algorithm

Note: It is expected to have submission in only one coding language with detailed explanation of the algorithms, limitation and conclusion in your documented repository

1. IC ICP Submission Guidelines:

2. ICP Submission is individual however, it can be completed as a Team during session.
3. If completed, should be presented to TA or Instructor before the completion of the class
4. Submission after the deadline is considered as late submission. (Check the late submission policy in the syllabus)
5. ICP Code with brief explanation should be pushed to GitHub.
6. Submit your screenshots as well to GitHub and documentation. The screenshot should have both the code and the output.
7. Submit a demo video 2-3 min showing your assignment with a voice over explaining your work if you are unable to complete ICP within the deadline due to genuine reason.
8. Provide the video submission link through the GitHub and submission form <https://forms.gle/xMAmr3zATrtMG5cX7>

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/>