CSEE 5590/490 Big Data Programming

Lesson Plan #6

ICP Feedback and Submission Link:

https://forms.gle/xMAmr3zATrtMG5cX7

Lesson Title: *Solr & Lucene*

Lesson Description: *Parallel Indexing: Solr and Lucene*

Lesson Overview:

Solr is highly reliable, scalable and fault tolerant, providing distributed indexing, replication and load-balanced querying, automated failover and recovery, centralized configuration and more. Solr powers the search and navigation features of many of the world's largest internet sites.

In Class Exercise:

Source: https://github.com/apache/lucene-solr/tree/master/solr/example

1. *Film Dataset*
   - Keyword matching
   - Wildcard matching
   - Proximity matching
   - Range searches
   - Fuzzy logic

   Dataset Description:
   https://umkc.box.com/s/6n4e9x73ruown2mxzfdfog1xfi9g8k50

   Dataset: https://umkc.box.com/s/z3lr0k0ejc6gz4zzf2239esxxz637sdm

2. *Books Dataset*
   Dataset:
   https://github.com/apache/lucene-solr/blob/master/solr/example/exampledocs/books.csv

ICP question: Execute any 5 queries on the given dataset

ICP Submission Guidelines:

1. ICP Submission is in pairs of three/four students.
2. Once completed, must be presented to TA or Instructor before the completion of the class
3. Submission after the deadline is considered as late submission. (Check the late submission policy in the syllabus)
4. ICP Code with brief explanation should be pushed to GitHub.
5. Submit your screenshots as well to GitHub and documentation. The screenshot should have both the code and the output.
6. 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.
7. 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/*