**CSEE 5590/490 Big Data Programming**

**Lesson Plan #6**

**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***

Dataset: <https://github.com/apache/lucene-solr/blob/master/solr/example/films/films.csv>

ICP question: Execute any 5 queries on the given dataset

1. ***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 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)
4. 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://forms.gle/wyFHFBL6LzM3xDGr5**)**

**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). Also submit your wiki to Turnitin where plagiarism should be less than 15%.
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

(ICP Submission Link#): https://forms.gle/wyFHFBL6LzM3xDGr5

***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/)