

## CPSC-51000: Introduction to Data Mining and Analytics

### Term Survey Paper

Throughout the term, you will need to work in an assigned group to produce a survey term paper on a topic within the field of Data Mining. The goal of a survey paper is to summarize latest research work on the given topic, group the approaches used by different researchers, outline the main themes of the current research and identify open problems. It should also describe the applicability of the current research to real-world problems.

For your survey paper, you will first identify an appropriate topic and select a few relevant papers to start researching. You should then do additional research by looking through the papers mentioned in those works as well as more recent papers that cite those papers. You should also use library resources and sites like Google Scholar to search for more papers on the topic. As you are researching papers, you should identify those which seem most relevant and interesting (look at abstracts; then intros and conclusions; skim through the rest of paper). At the end, you should collectively read and study at least 10 scholarly papers on your topic. However, throughout the process, you should go through many more papers. I expect that in the final version of your survey term paper, you will have at least 20 citations to quality papers. Your paper will be judged on the number and quality of sources, how you managed to group the approaches used by different researchers, outline the main themes of the current research and identify open problems, and by the quality of your writing.

#### **Timeline**

Week 1 (by Sunday, 11:59pm CST): Select topic, do some initial research, and describe it.

20% of overall term paper grade

Week 7 (by Wednesday, 11:59pm CST): Submit survey term paper draft

20% of overall term paper grade

Week 8 (by Sunday, 11:59pm CST): Submit final survey term paper draft

60% of overall term paper grade

#### **Week 1: Topic Choice**

For this first week, you should work in your group to pick a topic for the term survey paper, identify some initial scholarly papers on the topic (min. 3), and provide a short (1 page) description of the topic and your preliminary research. Some ideas for papers include:

- storage/processing systems (hadoop/spark/...)
- text mining approaches
- clustering large data sets
- rule mining (e.g. sequential rule mining), rule mining metrics
- machine learning (classification/regression) for very large data
- geographical data mining
- link analysis on very large data / link analysis metrics
- e-commerce analytics
- recommendation systems

Your report should be in IEEE styling, include the names of all group members, and have appropriate scholarly citations. Use the supplied template.

Good sources about data mining can be found here:

<https://www.omicsonline.org/data-mining-journals-conferences-list.php>

I recommend using Lewis University Library resources and Google Scholar to search for sources. See Week 7 for requirements of the survey term paper that

Grading:

0-2: Relevance of selected topic

0-2: Presentation (well written report, good use of figures and tables when appropriate, used references when appropriate, no spelling or grammar mistakes)

0-2: Following directions (submission format, page length limits, etc.)

### **Week 7: First Draft**

Submit the term paper draft by Wednesday, 11:59pm CST.

Grading:

0-5: Sources (number and quality)

0-5: Content (summary of research, how you managed to group the approaches used by different researchers, outline the main themes of the current research and identify open problems)

0-5: Presentation (quality of your writing, spelling/grammar)

0-5: Following directions

### **Week 8: Final Draft**

Submit the final version of term paper.

Grading

0-5: Sources (number and quality)

0-5: Content (summary of research, how you managed to group the approaches used by different researchers, outline the main themes of the current research and identify open problems)

0-5: Presentation (quality of your writing, spelling/grammar)

0-5: Following directions

The paper needs to be saved as a PDF and needs to adhere to the IEEE Conference Proceedings style format (US format). You can find a template for this format here:

[http://www.ieee.org/conferences\\_events/conferences/publishing/templates.html](http://www.ieee.org/conferences_events/conferences/publishing/templates.html)

### **What to Turn In**

A single PDF document containing your report.