COMP 6940: Big Data and Visual Analytics
Project Overview

1 Project Overview

This project will provide students with practical experience in handling, analyzing, and drawing insights from real-world datasets. Students will work in groups of four to tackle a non-trivial data science problem by applying appropriate algorithms, computational techniques, and analysis methods. Work done will be integrated into an application that effectively showcases the value of your findings, either through visual means or other impactful methods. **We have a zero-tolerance policy for plagiarism**; any instances discovered will lead to severe penalties, including a substantial deduction in your grade. Feel free to use any suitable programming languages and libraries to complete your projects, once all sources are documented.

2 Group Formation

Students are required to form groups of four. Each group is encouraged to bring together a diverse set of skills and backgrounds to effectively address the project's challenges. Should any student feel that the workload distribution within your group is uneven, we ask that you bring it to your lecturer or teaching assistant for a fair resolution. To ensure transparency and equity, each student will undergo a brief interview to assess their actual contribution, which will then inform their portion of the project grade.

3 Project Proposal

Due Date: March 20th

Each group must submit a one page project proposal outlining:

- The problem they intend to solve.
- The dataset(s) they plan to use, including a brief description.
- Preliminary ideas on the analysis methods and algorithms to be applied.
- Expected outcomes and deliverables.

The proposal will be presented to the class, and groups will receive feedback to refine their project ideas.

4 Progress Report

Due Date: April 03rd

This includes a draft of the final report formatted in LaTeX using the IEEE conference template and containing at least the following:

- A finalized problem statement.
- A detailed data description and source.
- An execution plan/proposed methodology, incorporating feedback from the proposal phase.
- A literature search relevant to the problem being addressed.

This draft report is crucial for ensuring that projects are on the right track and will account for another 10% of the project grade.

5 Final Presentation

Due Date: May 01st

Groups will deliver a 15-20 minute presentation on their project, showcasing their analysis, results, and insights. The presentation should demonstrate a clear narrative, the methodology applied, and the conclusions drawn. This will be followed by a Q&A session.

6 Final Report and Code

Due Date: May 05th (11:59 PM) - **No Exceptions!**

The final report must be at least 10 pages long and formatted in LaTeX using the IEEE conference template. It should detail the problem statement, methodology, analysis, results, and conclusions. The report should also reflect on the project process, challenges encountered, and how they were overcome. Along with the report, all code developed during the project must be submitted. This combined submission accounts for 20% of the project grade.

7 Grading Criteria

Overall Weight: 50%

- **Project Proposal:** 10%
- **Progress Report:** 10%
- **Presentation:** 10%
- **Final Report and Code:** 20%

8 Possible Publication

Upon the successful conclusion of this project and the subsequent allocation of marks, we will review of all submitted reports to identify any with the potential for publication. Should your work be selected, we will notify you and offer our support in refining your submission for publication. Please note, this collaborative process will commence only after the completion of the course.