DS 340W: Applied Data Sciences

Team-based Term Project Description

55% of the course grade will be based on a team-based term project and writing. The purpose of this project assignment is to facilitate hands-on training in creating a data science solution to a real-world problem. Students are expected to use the provided state-of-the-art supercomputers (the Bridges CPU cluster and/or GPU cluster) for their experiments using real-world datasets. Students are to work with the instructors at the beginning of the semester to determine a problem to work on. *If you cannot find a suitable problem, the instructors will supply a problem for the students to tackle.* Throughout the semester, students will receive only high-level guidances from the instructor and TAs, and are expected to tackle the implementation issues on their own, possibly with the help of online resources.

The students are expected to go through an entire data science project pipeline, including understanding the real-world problem, formulating the problem as a data science problem, collecting data, formatting/cleaning data, developing a solution, experimentation, analyzing results, and articulating and defending findings.

Please follow the steps specified here.

1) Draft a one-page to 1.5-page project proposal (using the IEEE Journal template on Overleaf) addressing the following issues: purpose of the data science project, who will benefit from it, what are the technical challenges you foresee, the data science solution framework to be developed, and implementation plan with milestones for Week 6 and Week 9. **If needed, discuss with the Dr. Wang to finalize the project idea soon.** The project proposal is due on **February 7 at 11:59PM onto Canvas dropbox**.
2) For feedback on progress, meet with Dr. Wang in his office hour (via Zoom) during the class time **(see class schedule)**. Detailed schedule will be made available later.

Evaluation Criteria (35+20 points):

The project will be evaluated based on the applicability, correctness, design, and functionality. The final report, in a format of a scholarly article, will count for 20 points.

3 Proposal (one page)
5 Progress after Week 6
7 Progress after Week 9
20 Final report as a scholarly manuscript
20 Final in-class presentation and demo of the functionalities (peer review)