CS333 - Project

Project Details:
The project should be a significant data analytics project that requires substantial effort and breadth of knowledge. It may be an applied project, a research project, or some combination of the two. Projects in data analytics will be in groups of at most 4 students.

Project Proposal
The project proposal should be completed near the beginning of the class to allow sufficient time for completion. Ideally, the student would enter the course with some concrete ideas of projects that they would like to pursue. The instructor will help identify potential projects for students who do not already have concrete project ideas. The written project proposal must be typed and should be approximately 2-3 pages long. The purpose of the proposal is to make sure that your project is reasonable and enables the instructor to provide useful feedback.

The project proposal should at minimum cover the following items:

1. Preliminary title and student name(s)
2. Abstract: Like the abstract that will ultimately appear in your paper. It should be one paragraph long and provide a motivation for the project, a high-level description of the project, and the project's main expected outcomes.
3. Project Description: A description of the project and the main steps necessary to complete the project. Below are some of the items that will likely be covered for all capstone projects:

- What data sets do you plan to use? If you must do significant work to get the data or convert it into the proper format, then describe the process and approximate effort required. If you plan to collect your own data (e.g., via a survey), discuss the collection process. Describe the total amount of data that will be available and justify that it is sufficient for the project.
- What is the main data mining/data analytics task? The description does not have to be long, but it must be clear and well-defined.
- What learning tools/toolkits do you plan to use (e.g., WEKA, Python Scikit) and what data processing methods and model induction algorithms do you plan to explore?
- How will you assess/evaluate your results and how will you determine the utility of these results? Are there other papers/results that you can compare them against, or is this a novel problem?
- What related work exists? Try to provide a high-level description of the major related areas and related papers that fall in each of these areas. Minimally the proposal should include at least 4 related papers (the final project should include more).

4. Timeline: The proposal should break down your project into major and minor work items and provide an estimate of the amount of effort required for each; a tentative timeline should then be assembled. (gant chart)

Project Paper

The project and its results should be documented in a professional manner. The instructor will provide IEEE Conference Paper Templates (both use two-column formatting) and the student is expected to use these for the paper. Different formats can be used with prior agreement from the instructor. The main body of the paper is expected to be between 6 and 10 pages (longer papers are fine but should be discussed with the instructor). Your paper should generally adhere to the following outline, which is typical for a conference paper, but minor variations are acceptable:

1. Abstract: summarizes the paper and the goals of the work (required)
2. Introduction: Introduces the project and what you are trying to do. May include some background.
3. Background: Depending on the project, you may want a separate background section, depending on how much background you want to include. For example, it may provide domain information for the domain that you are studying.
4. Experiments: Describes the experiments and the experimental setup. Will describe the data sets, the evaluation metrics, the data mining tools used, and any other details related to the experiments.
5. Results: Includes the experiment results (which are typically not included in the experiments section). A discussion of the results may be included, or they could be included in a separate discussion section, which follows the results. For now, we will assume no separate discussion section.
6. Related Work: A brief description of related work, with citations to relevant papers. Research oriented papers will typically have more references, but all papers must contain at least 6 references.
7. Conclusion: Provide your conclusion. For example, comment on the quality of your results. You may also want to include some material on future work, whether you intend to do such work.

Oral Presentation:

Each project must be presented to the class via an oral presentation. Presentations should generally take 20-30 minutes. Visual aids (e.g., graphical results) should be provided.

Target Dates:

Project Proposal: TBA

Oral Presentation: TBA

Project Paper: TBA