CS 6220 Data Mining - Project Proposal
Due: November 1, 2023(100 points)

NAME(S)
PROJECT GIT REPOSITORY

1 Proposed Objective

The objective is a concise statement for the problem that you are solving.

2 Motivation

One common framework when considering problems to solve is the Heilmeier Catechism. In this section, try to think through the catechism to properly motivate your problem and describe why it is important. Make sure you mention what impact your work will have. More than verifying your hypothesis, what will your solution enable you to do? What are the consequences of solving this problem?

3 Background

This section should also describe any literature that you plan on reviewing or survey. These references can be related to both the applications of the work as well the technical approaches to be taken. Emphasize why they are appropriate and relevant.

If you're choosing a project besides Greenstand, make sure you describe the specifics of your data and what information that you believe would be most useful from it. If you're using the Greenstand data, articulate what features and data you plan on leveraging.

4 Proposed Approach

In your approach, discuss the objective that you are attempting to satisfy and your approach to optimizing for it. There are many factors, and learning algorithms may be a small portion of your processing. In your proposal, include how you're considering the scale of data, the noise

that you might face, the imbalance of the feature values and labels. Specifically (and briefly), address the following bullet points.

- what pre-processing you'll need to do

- how you intend on remediating class imbalance

- what learning algorithms you intend to try

- how you will validate and evaluate your results

You can work on this project on your own, or you may work in groups of four at most.

5 Submission Guidelines

Upload your PDF file to Gradescope as a *single* submission with all members of the team. On the top of the document, include the teammates and project repository. *Try to keep your proposal below three pages.*