NORTHEASTERN UNIVERSITY, KHOURY COLLEGE OF COMPUTER SCIENCE

CS 6220 Data Mining — Assignment 6
Due: March 29, 2023(100 points)

YOUR NAME
YOUR GIT USERNAME
YOUR E-MAIL

Your assignment for this week will be to send us a 2-page project proposal. The Google Docs version of this proposal is available here. You can choose the format you want to use, but make sure to include the following information to your document:

1. What problem are you going to be tackling on your project?

2. Why is that an interesting/useful application of data mining?

3. What models/techniques (clustering/classification/etc.) are you envisioning to apply?

4. Where are you going to get the data?

To complete your proposal submission, create your project's repository in your own Github namespace, and upload your provide your URL in Gradescope. For example, my Git handle is `kni-neu`, and my project repository is:

```
https://github.com/kni-neu/project
```

You can use my template or any other template that you might find appropriate. (Or, it is acceptable to use no template at all.) An example of a set of organized proposal sections is shown below.

1 Problem

Make sure you make clear what problem are you going to solve in as concise and unambiguous manner. This section is typically 2-3 sentences.

2 Background

This is where you tell us why this solving this problem is important. What will people be able to do once you've solved this problem? How could it conceivably help people? What makes this application of data mining useful?

3 Approach

Here, you can start detailing some specifics. Be sure to cover:

- Where are you going to get data?

- What data mining techniques will you use?