1. What are the names and NetIDs of all your team members? Who is the captain? The captain will have more administrative duties than team members.

Name & Role: Saurav Sharma (captain)
NetID: saurav4

2. What topic have you chosen? Why is it a problem? How does it relate to the theme and to the class?

Category: Free Topic
Topic: Dota 2 Hero State Analyzer
Context: Dota 2 is a complex video game with over 100 heroes to choose from. Changes in the game occur at varied cadences through patches released by Valve (the company responsible for Dota 2).
Since there are so many heroes in the everchanging game, it can be difficult to identify heroes which have either been consistently buffed or nerfed over the last few patches.
My goal with this project is to build an application which can let me know how many, and which, buffs/nerfs a hero has received in the last N patches.

This relates to the theme of this class because I will be parsing text information, and feeding it into a classification model that I will train to identify whether a change is a buff or a nerf.

3. Briefly describe any datasets, algorithms or techniques you plan to use

Patches will be scraped from https://www.dota2.com/patches/
A database must be built which contains all heroes, and all changes (along with the patch the change is associated with).

I will then manually label 2-3 patches worth of data and mark changes as buffs, nerfs, or reworked (sometimes abilities are fully changed and cannot be remarked as a buff or nerf).
Using this labeled data, I will train a classification model to identify whether a change is a buff, nerf, or rework.
With this model, I hope to build a basic python application that can return how many nerfs and buffs a hero has received in the last N patches.

4. How will you demonstrate that your approach will work as expected?

I intend to include a confusion matrix in my report which can showcase the validity of the approach.

5. Which programming language do you plan to use?

Python

6. Please justify that the workload of your topic is at least 20*N hours, N being the total number of students in your team. You may list the main tasks to be completed, and the estimated time cost for each task.

This is a single member team. I have never done such a task and will outline what I believe is involved below. I sincerely believe that the totality of work will exceed 20 hours.

a. Write a webscraper which will scrape https://www.dota2.com/patches/
b. Create and populate a database with heroes and changes
c. Divide scraped patches into training and test datasets.
d. Label training data.
e. Create a classification model with training data.
f. Test the model and create a confusion matrix for evaluation purposes.
g. Write a text-based Python application which uses the model and answers the question of “How many (and which) buffs & nerfs has X hero received in the last N patches”.