CS 410 Project Proposal

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.

    - Akash Patel - apate400 - Captain
    - Rushil Thakkar - rthakk20
    - Ebubechukwu (George) Obetta

2. What is your free topic? Please give a detailed description. What is the task? Why is it important or interesting? What is your planned approach? What tools, systems or datasets are involved? What is the expected outcome? How are you going to evaluate your work?

    - Stock market sentimental analysis to predict stock appreciation and depreciation based on tweets on Twitter for a specific stock. This will be done by training a BERT model with test data with past tweets that are labeled as positive or negative based on the sentiment.. The tweets that will be run through the sentiment analysis will be gathered through the Twitter API. The overall goal of our project is to have a user type the name of a stock, and then we gather the current tweets for it and let them know if it has a positive or negative sentiment and if the stock will go up or down based on that.

    - We plan to use the stock market tweets dataset from the kaggle website here: https://www.kaggle.com/utkarshxy/stock-markettweets-lexicon-data.

    - We plan to use the following Twitter api endpoint to search for tweets about a stock: https://api.twitter.com/2/tweets/search/recent?query=

    - Once the model is trained then we will use it to classify and count the number of positive and negative tweets for a stock for a specific period of time. Then we will predict if the stock will appreciate or depreciate the next day based on this data.

    - To evaluate our work, we can predict the sentiment for a stock and see if the prediction we gave stands true or false.

3. Which programming language do you plan to use?

    - We plan to use HTML, CSS, and JavaScript for the front end and python for the backend and running the sentiment analysis models.

4. 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.
    - Learn BERT model & implement it - 30 hrs
    - Front End - 15 hrs
    - API Development & Integration - 10 hrs
    - Debugging & helping teammates with the problems in their respective tasks - 5 hrs