CS410 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.**
   a. Matthew Jin, mjin11
   b. I am the sole member of this group, and as a result, I am also the captain.
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?**
   a. Everyday, I watch NBA basketball. And as a fan of basketball, I enjoy imagining what players would fit well together. I almost act as if I am the General Manager (GM) of my favorite team, thinking about what trades I need to make this team the best it can be. This project aims to automate the GM experience. Given the NBA salary cap and current team, my model will suggest and rank available and valid trades to make so the team can be the best it can be. For example, given a team with a lack of 3pt shooters, my model will suggest trades that acquire said shooters and still keep the total salary under the cap. Evidently, this model will know that it can't just recommend trades for the best and highest-paid players in the league. This model can be pivotal for not only basketball trades but also trades in other sports (baseball, football, etc.).
   b. The main dataset I plan to use is from https://www.basketball-reference.com/. The basketball reference API allows me to obtain player stats, team stats, and financial information in the NBA. I plan on using data processing and statistical analysis to mine and understand player stats. I plan on using a vector space model to project players into a vector space. Similar players will be closer in this vector space. What the dimensions represent will be decided from the analysis I do on the player stats and player contracts.
   c. The expected outcome of this project is a simple user interface that takes in a team of 15 players and outputs the top ranked trades that the model thinks will help this team win a championship. The evaluation of my work will be difficult as it will be hard to simulate an actual NBA season. However, there are games (NBA 2K) and online websites that allow you to simulate a season. I would like to take advantage of these video game simulations and also manual judgements through the use of online polls in social media, asking the people's thoughts on if these trades will improve the team.
3. **Which programming language do you plan to use?**
   a. I plan to use Python. The Python libraries that I will most likely use are scikit-learn and PyTorch.
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.**
   a. Understanding and using the API to grab data from basketball-reference.com (3 hours)

b. Analyzing player stats and exploring data processing and statistical analysis through scikit-learn to represent player stats in a fixed number of dimensions (8 hours)
c. Implementing core algorithm and function to find trades by imagining a new player and matching that fake player to a player that already exists (6 hours)
d. Implementing a simple user interface in Python (1 hours)
e. Testing the model through video game simulation and polling on social media (3 hours)