DRP Reflection: NBA Analytics and Machine Learning

Mentee: Haoquan Fang

Mentor: Andrea Boskovic and Harshil Desai

During the Spring Quarter of 2023, it is my pleasure to know and join the Directed Reading Program held by the Department of Statistics at UW. Through this valuable opportunity, I finished my research project under the general topic of "NBA Analytics and Machine Learning". I also had a great time working under the instructions of my wonderful mentors Andrea Boskovic and Harshil Desai. I feel appreciated that they guided me to conquer the difficulty and directed me to my success in this program.

To be more specific, I would call my project "NBA Game Wins Prediction Based On Machine Learning Algorithms," which aims to utilize machine learning techniques to predict the number of game wins a team will have in a particular season using its performance statistics in the previous season. Through the project, I started from scratch and finished a complete machine learning project pipeline: collecting and preprocessing the data, feature selections, and developing various machine learning models to forecast game wins.

This project provided me with a valuable learning experience on multiple fronts. From a technical standpoint, I developed a deeper understanding of data collection and preprocessing techniques. For example, when I was aggregating the performance statistics of a team in one season, I learned that taking the average of the performance statistics of all players in the team will be skewed by some outliers, and that was why I chose to take the sum of those statistics. I also gain more theoretical knowledge about the purpose of feature selection. I learned some basic optimization concepts that fewer features will make training loss higher and testing loss smaller, and vice versa. From a practical standpoint, I became more experienced in dealing with different practical challenges. I faced challenges when dealing with missing game data and strange training loss, but through extensive research and consultation with my mentors, I was able to implement effective strategies to handle these issues.

I am immensely grateful for the guidance and mentorship provided by Andrea Boskovic and Harshil Desai throughout this project. Their deep knowledge of machine learning theory and application significantly enhanced my understanding of the subject matter. Their consistent availability and willingness to answer my questions and provide feedback played a crucial role in the success of this project. In our weekly meeting, I also had the special opportunity to communicate with another team member, exchanging ideas and insights, which fostered a dynamic and stimulating learning environment.

As a result of the collective efforts, I was able to develop various machine learning models that try to forecast the game wins of a team based on its historical data. The feature data and predictive models I created would provide valuable insights into the field of sports prediction. Though it was not perfect, my project is still a good start in my learning career. I definitely value this opportunity and experience in this program, and I was willing to make some future improvements to the project if possible.