USING MACHINE LEARNING TO HELP BASKETBALL TEAMS WIN MORE REGULAR SEASON GAMES

Mandla Senzanje, Rajeev Bukralia. Department of Computer Information Science, Minnesota State University, Mankato, MN 56001

With the game of basketball continuously evolving, teams are always finding ways to stay ahead of the competition, by winning more games and remaining a profitable/relevant franchise. This research tackles two questions: how do various basketball stats contribute to a team getting a win or loss (target variable)? How does a team’s performance affect attendance and Wikipedia popularity? To answer the research problem, regular season games data from the NBA website (2012 – 2018), and team Wikipedia page view data was collected. Initial steps will involve data pruning and data manipulation (coding categorical features such as team name, conference, and location into binary). Splitting the dataset will follow, which will be necessary for the feature selection portion. Various machine learning models will be trained, and whichever model has the best accuracy scores with the test data, will be used for feature importance graphs and used to predict wins and losses. Next step will be to analyzing the relationship between team Wikipedia popularity with the team’s performance and attendance. This research can provide insights to basketball teams on two aspects. Based on a team’s current record and stats, they can know what stats they need to improve upon. Having better knowledge on the team’s performance they can know what they need to improve their Wikipedia popularity and attendance. This research can be applied towards different types of sports to improve a team’s winning rates and can recommend fantasy sports customers on what type of players they should get.