Jonathan Greenberg

Report #5

3/1/2020

By the time we meet again on Thursday at 1PM I will have worked on the following:

1. In EDA I will further create categories out of the notes column for injuries. Some ideas for categories are: body part, injury, severity of injury, time out for (DTD, indefinitely)

Also add a binary 0/1 column for whether or not the player was returning to the lineup, or he was sitting out of the lineup due to the injury or whatever the event was.

2. Create more visualizations, look into multidimensional parallel line visualization for types of injuries or players.

3. Send e-mail to the capstone project advisor with a precise budget so we can request funding. Mention how the data would help improve your project.

Also try to get game statistics for pre-injury, like how many minutes played, how well they were shooting, etc.

If I am unable to get the schedule/rest day data, I can add an alternative question: Can we predict based on physical attributes what kind of injuries players are more likely to get, thereby injury prevention can be more tailored to each player?

4. Think about ML models. Prediction? Classification?

* Using the newly generated features for types of injuries, can I predict a given feature.
* Using neural networks, decision trees, etc.
* Clustering analysis for the types of injuries, or the players, or the teams (i.e. which teams have the most similar injuries)