Jonathan Greenberg

3/22/2020

Project Report #7

I am still waiting on crucial data to be sponsored by YU that will help in my predictions

Until then, I am working with what I have in analyzing all of the injuries in the NBA in the last 10 years.

I have successfully created three new columns in my dataframe called: “body”, “injury”, and “severity” which extracted detailed information about each injury log and categorized them into these three columns. I am now looking into ways that these new columns can provide more insights.

All of my work thus far can be seen in this colab:

<https://colab.research.google.com/drive/1InZ1b1YDH97RYYncTBs5TKvWcNoX9PPr>

There isn’t much modeling I can do without the historical game and scheduling data, but I am looking for other ways to do clustering with NLP and make word embeddings to see which players and teams have the most similar injury types.