Data Memo

Aidan Baker

4/10/2022

Overview of data set

I am going to doing my Machine Learning project on the UFC. The dataset that i will be using entails names of fighters, their win loss record, height, weight, reach, as well as key fight statistics such as strikes per minute, strikes landed per minute, accuracy, and takedown accuary. In total there is about 15 or 16 predictors that will be going into the model. Currently most of this data is avialable online publicly, additionally I have a connection who works for Endeavor, which is the company that manages the data analytics of the UFC. I am going to be meeting with him on 4/12 to discuss if he has any additional key data that could be beneifcal for my project. There is some missing data but not much, for some of the fighters there are a few unknowns such as a specific fighters reach or date of birth. I think that it will not be hard to work around this data.

Overview of research questions

I am interesting in predicting who would win between two fighters based on past wins as well as performance in key fight statistics. I believe that this question is best answered through a classification approach since it is a win or loss perdction and not a continuous value that would typical be solved through a regressional approach. I think that the key predictors for my model are going to be SLPM (strikes landed per minute) as well as Takedown average and accuracy. As a UFC fan myself, I believe that these are key statistics that usually determine outcomes of fights. I think that the goal of my model is predictive as it is predicting the outcome of a fight between two fighters.

Proposed Project Timeline

I plan on having my data set loaded by the end of the week. I hope that I am able to obtain additional data by the middle of the week and hope to encorporate that data into R and starting the process of the project by this weekend beginning the exploratory data analysis. As for the rest of the quarter, i plan on keeping up to date and implementing new methods and techniques as we do them in lecture to stay on top of everything.

Questions or Concerns

As of right now i do not think there are any problems or difficult aspects of the project. Although i do suspect that down the road i might encounter something that needs to be independently researched and taught since it might not be a topic of the course. I do not have any specific questions for you at the moment but will make sure to ask if any arise.