

Alex Ojemann
2/27/2023
Pitch Feedback

Neural Net Ninjas:

Problem space: I don't understand how predicting a player's age based on other features allows you to achieve your goals of identifying talented young players and deciding which players a club should invest in. Is it desirable to have a lower predicted age than actual age or a higher one? How does predicting age help to identify talented young players when clubs already know their age? Athletes usually have a career arc where they're at their best in the middle and at their worst at the beginning and end, so it could be hard to classify between young and old players. A better approach in my opinion would be to use age as a feature along with the others you mentioned to try to predict an overall value metric (like WAR in baseball) for the rest of a player's career. This would better help teams judge how much a player is worth paying.

Dataset: The data is well explained and clear for the most part. As long as you scale the features the differing variable scales won't be an issue. I just wonder about how you will prune the features, 110 features with over 15,000 entries seems like it will be way too much data.

Methods: I wonder how you will evaluate if you're successful in that I don't know how better age prediction translates to better player evaluation.

Misc: Nothing else

Team KGS:

Problem space: I think the problem space was explained well.

Dataset: I noticed you guys have 29 features for most players but only 6 for goalkeepers. I'm not sure how you will be able to put those in the same model. I think it may be worth doing separate models for goalkeepers because the numbers of features and the value of the position are so drastically different. I also wonder if there are any bias in the Fifa ratings. I don't play Fifa but I know that in NBA 2K many popular players or players in big markets tend to have higher ratings than their statistics would suggest. It may be worth separating the Fifa ratings and the stats into different models and comparing their performance.

Approach: I like that you're using four different methods to see which is most predictive. The only thing I'd note is that I saw in the video you said that you wanted to use KNN to classify win loss or draw whereas with the Naive Bayes you were going to output win probabilities, you'll have to figure out what the optimal probability thresholds are to use for each of the win loss and draw categories if you want to compare the different models.

Misc: Nothing else