

The experiments result is 90% accuracy for all tested player that is not outliers. Instead, if we use the average salary as the prediction, the accuracy is only 80%. By using this method, the accuracy increased a lot. So the claims are supported by the experimental result.

We have tried to predict the salary by using the film rating pattern taught in class, but it gives us a poor result, the predicted value are always around 15, so that result doesn't make any sense. We also tried to predict the salary by the ratio of positive tweets and negative tweets, but it seems that the salary are more likely to go with how many people talked this player in twitter, instead of how much this player was liked/disliked by people.

This experiment need to retrieve data from the database, and the database file is 14G, include 1.3M tweets, so we cannot upload to github. However, the ipython notebook will represent all the steps that we did and all the data that we collected. The code is runnable and executable.