Homework 4

1. What I found interesting is the four weighing schemes. Each of them seems reasonable and produces a different result, but there is no best method. In your paper, you included all four weighing schemes and their outcomes, so I wonder if we can submit as many (or at least a lot) predictions as possible, and the accuracy of prediction tends to improve as we submit more and more. If we don't know the best method, we may just include some randomness in each scheme and duplicate many predictions, and then we can predict better.

Also, about the difference in Colley and Massey methods, both methods are reasonable, and the theory part is not that hard, but I don't know why their outcomes differ a lot. More specifically, the four weighing schemes for the Colley method have a large percentage of variance, but the schemes for the Massey method don't. The four percentages for the Colley method are 62, 65, 65, 97, and the four percentages for the Massey method are 88, 79, 81, 79.

I enjoy the application part of the paper more than the theoretical part because I prefer the application, which is visual and exciting, than theories.

2. In my dataset, there is data for NFL 2020 Player Stats for Rushing. The data is sorted by Pass Yards descend. It includes players' names, Pass Yds, Att, TD, Lng, and many other aspects (I just copy down the titles but I don't know what they mean, I don't watch NFL). I used Google Sheet to scrape the dataset.

3. I will do this homework in the next session.