Sam Tracht

CMSE 381 Final Project

1 Project: Nate Silver ELO Score Reproduction

1.1Methodology

For this reproduction we will be using a time series model. From the dataset, NBA-ELO, we will try and use the games scores from 1946 to 2015 to reproduce Nate Silver's ELO rating system. This model will be based on the formulas shown in "What is an ELO rating?" (Mittal), with adaptations for the difference in style of NBA ELO.

1.2 Reproduction

From "What is an ELO Rating" by Mittal and "How We Calculate NBA ELO Ratings" we got an initial time series formula to update ELO. The only initial change from the chess ELO updating equation and ours is that the optimal K is found to be 20 as well as updating ELO on new seasons toward 1505, this is to give a slight reset to teams, but three quarters is kept because NBA teams are consistent season to season (Silver). Unlike chess, basketball has outside factors that can play into how well a team performs, one of them being whether they are home or away. Home-court advantage is set as equivalent to 100 ELO points (Silver). In our time series after getting the previous ELO we will then add 100 to it before getting the expected scores. Along with the update for home court advantage, K is not constant like in chess, Nate's model considers the margin of victory as well as the ELO difference between two teams (Ergo Sum). Initially I set the model to a constant K and did not factor in for home court advantage, the correlation for between my model and Nate's model for the Los Angeles Lakers was .976. The overall correlation between every team was un-weighted .88 and weighted .95 (teams with more games played will weigh more into the correlation than that of teams who played less). After adjusting the model parameters and rerunning the time series the model increased in accuracy. The correlation for my adjusted model and Nate Silver's ELO scores for the Los Angeles Lakers increase to .987. Along with an increase in the Lakers model, the overall score for all the teams increased, the un-weighted correlation among all teams increased to .9288 and the weighted correlation increased to .9775. Additionally, the model posed to have a slightly higher correlation between final ELO ratings increasing from .976 in to .989 and decreasing RSS from 146808 in the initial model to 119176 in the new model shown in Figure 1.

Figure 1: Los Angeles Lakers ELO Score Comparison Over Time: Tuned Model

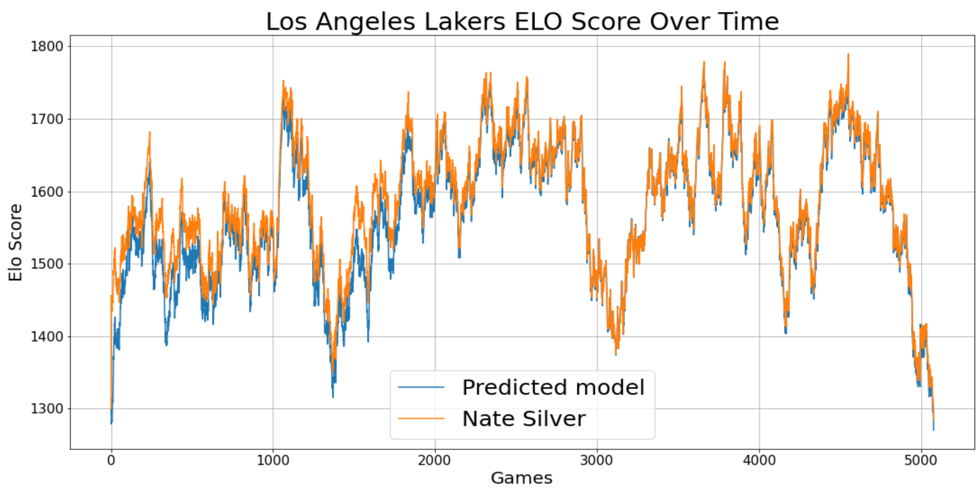
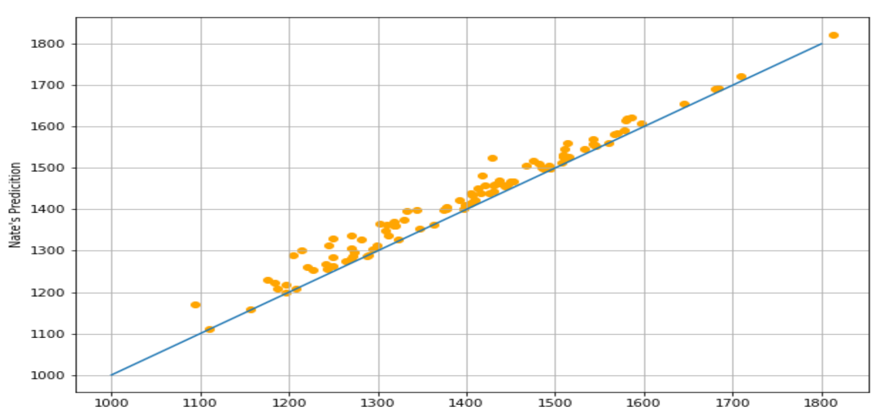


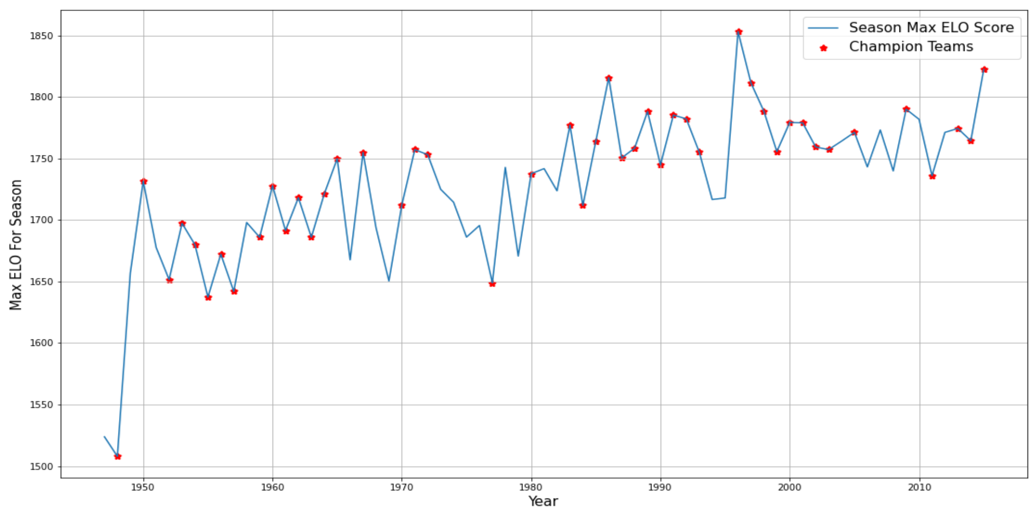
Figure 2: Teams Final Standings on My Model vs Nate Silvers: Tuned Model



1.3 Additional Analysis

While this model was used to try and continually update ELO scores it did not tell how ELO can be used as a predictor. We analyzed the maximum ELO score in an NBA season and determined if that score was produced by the champion for that season. The maximum ELO score of a season was produced by the team who won the championship 67% of the time as shown in Figure 5.

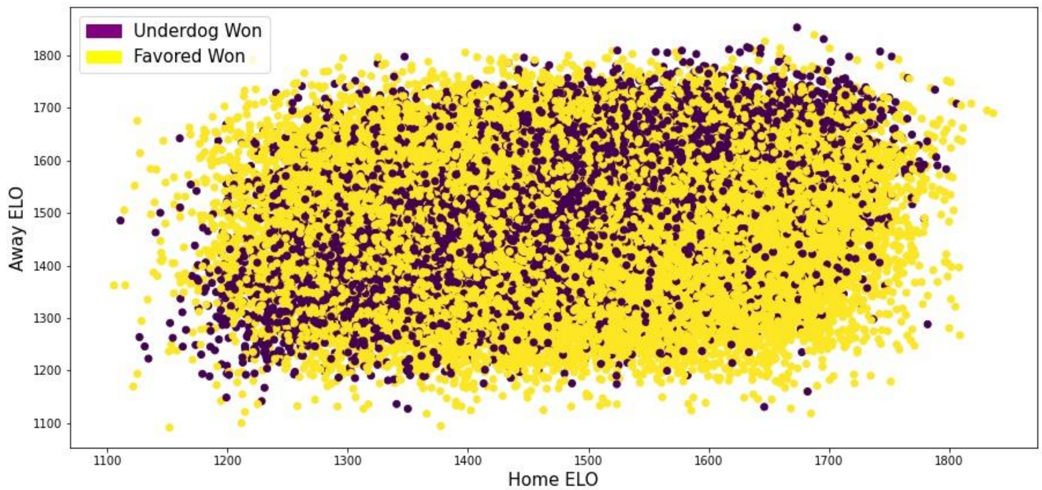
Figure 3 Max ELO Score for each season



1.3.1 Classifiers

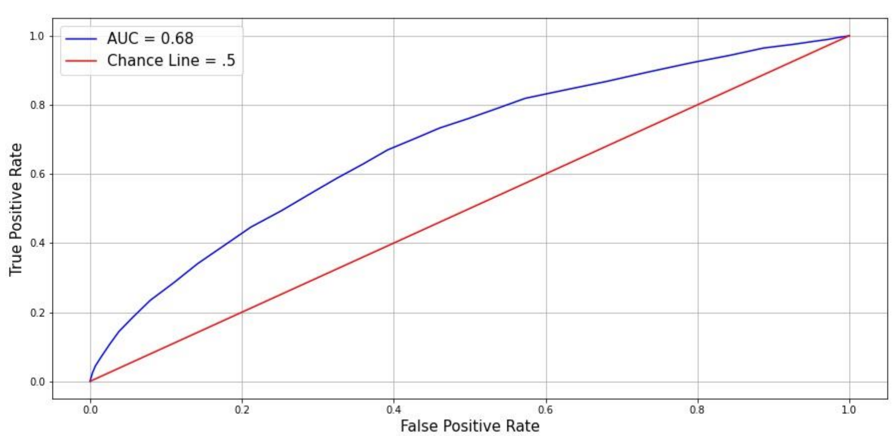
The team with the higher ELO won 64.33% of the games from 1947-2015 with Nates Scoring system, besting mine which scored 64.2%. From this, I was interested in seeing if a classifier such as KNN or SVM could better predict whether the favored team would win or not. Using the initial game ELO and Score I classified each game based on if the team with the greater ELO won the game or not (In the case of ties in ELO score favor was given to the home team). The general trend, shown in Figure 4, is that underdogs won more when the ELO scores were closer and that favored teams won a majority when they had a large margin between the ELO scores, while also still winning close matchups.

Figure 4: Initial Games Classification



From this data I created a KNN classifier with a K value of 5, and it predicted less correct, 63.33%, than if you were to just choose the team with the higher ELO from Nate's Model. From this I repeated this process on K values from 1 to 100 and found the optimal K value to be 46, having an accuracy score of 67.8%, finally outperforming Nate Silver. Although this was the best classification, the accuracy scores only fluctuated ~2% from K=10 to K=100. Figure 5 shows the actual performance of the model compare to just flipping a coin and picking the winner.

Figure 5: AUC Curve for KNN model



1.4 Conclusion

This model reproduction proved to be difficult in precisely replicating but fairly easy to get a general replica given the equations from the three different sources. Overall, my models performed fairly similar to Nates. One place for improvement is analyzing the under prediction of my model. In figures 1 and 2 you can see that I was consistently underpredicting the ELO score for the teams. Along with that, the models K value could be tuned using more variables such as win streaks or injuries/trades. Additionally, ELO seems to be a generally good predictor for wins, considering the only variables that play into a team's ELO score are points score (at least for the updated model). With that I believe adding variables such as team statistics this model's accuracy could improve vastly. I believe the same to be true for the KNN model in this case, I think that using additional information to classify the results would benefit the model beyond what and ELO score alone could.

1.5 REFERENCES

1. Mittal, Raghav. "What Is an ELO Rating?" Medium, Purple Theory, 6 Nov. 2020, medium.com/purple-theory/what-is-elo-rating-c4eb7a9061e0.
2. Silver, Nate. "CARMELO NBA Player Projections." FiveThirtyEight, 8 July 2019, projects.fivethirtyeight.com/carmelo/.
3. Silver, Nate. "How We Calculate NBA Elo Ratings." *FiveThirtyEight*, FiveThirtyEight, 21 May 2015, fivethirtyeight.com/features/how-we-calculate-nba-elo-ratings/.
4. "Replicating Nate Silver's NBA Elo Algorithm." *Ergo Sum*, 30 Apr. 2018, www.ergosum.co/nate-silvers-nba-elo-algorithm/.