THE COLLEGE OF WOOSTER

Predicting Formula 1 Races

Using Ordered Logistic Regression and Elo Ratings

Luke Pritchard, Statistical & Data Sciences, Advised by Dr. Luri

Introduction

This project investigates the Formula One World-Championship in the context of predicting race results. Formula One is an incredibly popular sport that takes place on 5 continents around the globe that has several billion dollars in revenue per year. Positions in each race can be a massive difference in terms of prize andsponsorship money/exposure. To predict these races, **Ordinal Logistic Regression** and the **Elo Ratings** method are used, combined, and compared. This research seeks to find out if Elo works well in this area, and if combining it with data from Qualifying sesions that take place before the race can be useful.

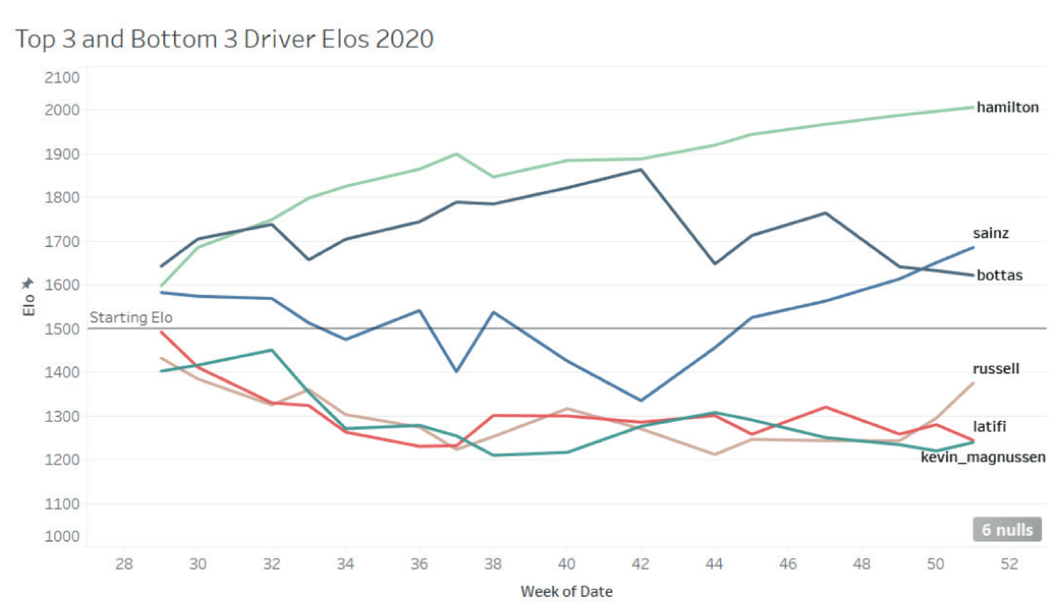
What are Elo Ratings?

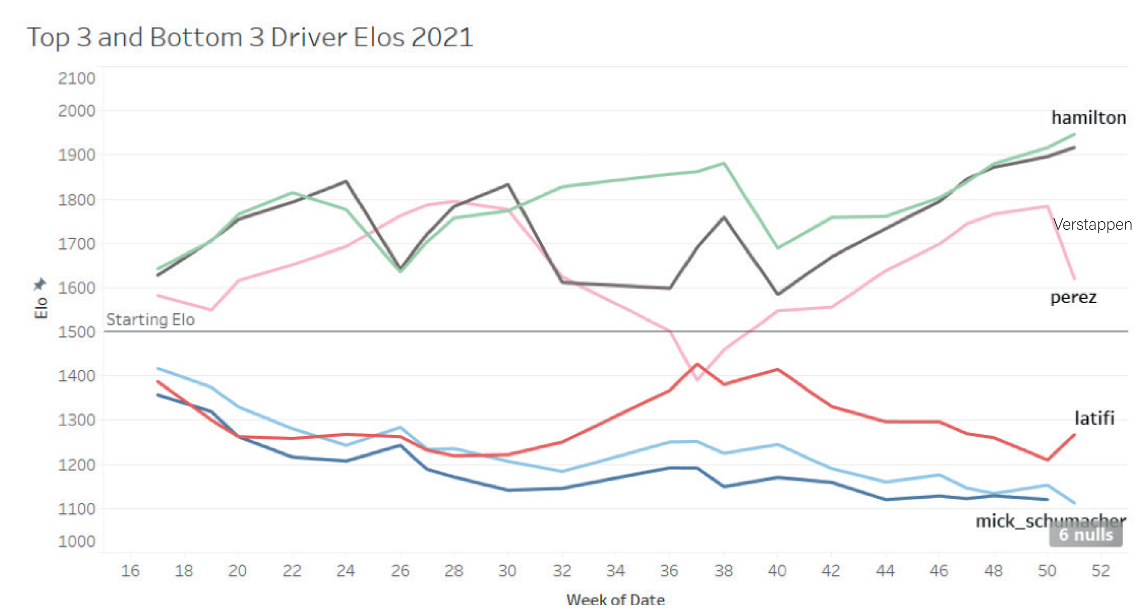
- Elo is a rating system designed to give the relative skills of players in competition. It was invented for chess by a man named Arpad Elo, a Hungarian-American physics professor who was born in 1903.
- After every match, points are exchanged between players. The theory behind the system states that after enough games, the true skill rating of a comepetior is revealed.
- The central assumption of Elo is that the performance of each player in every game is a random variable that follows a normal distribution over time.

$$E_A = \frac{1}{1 + 10^{(R_B - R_A)/400}} \qquad E_B = \frac{1}{1 + 10^{(R_A - R_B)/400}}$$

$$R'_A = R_A + K(S_A - E_A)$$

Elo Ratings Over 2020 and 2021 Seasons



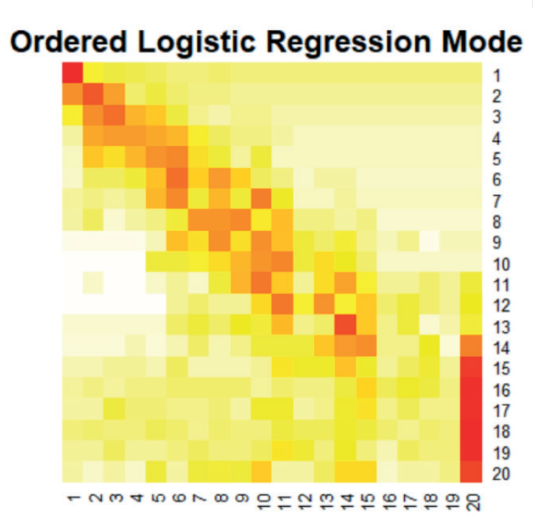


- Using a package in R, I created Elo ratings for every driver updated after each race for the 2017-2021 seasons of Formula One racing.
- To calculate this, each driver was put in a matchup against each other driver; this is different than most implementations of Elo where it is a 1 versus 1 matchup.

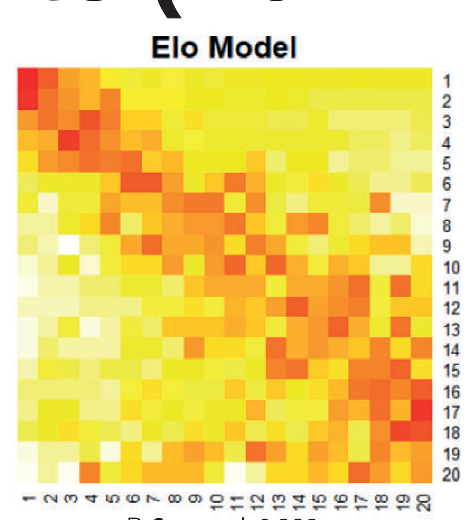
Creating the Models

- First, I made a logistic regression model without elo. It predicts finish position of a given driver based on percentage of pole time, starting grid position, the team they drive for, and their average finish in the previous races that season.
- Next, I looked at how accurate the Elo ratings were to predict each race. In theory, the order of the ratings is the most likely order of finish.
- Last, I made a model combining the first model with the Elo ratings to see if, working together, they could create an optimal result.

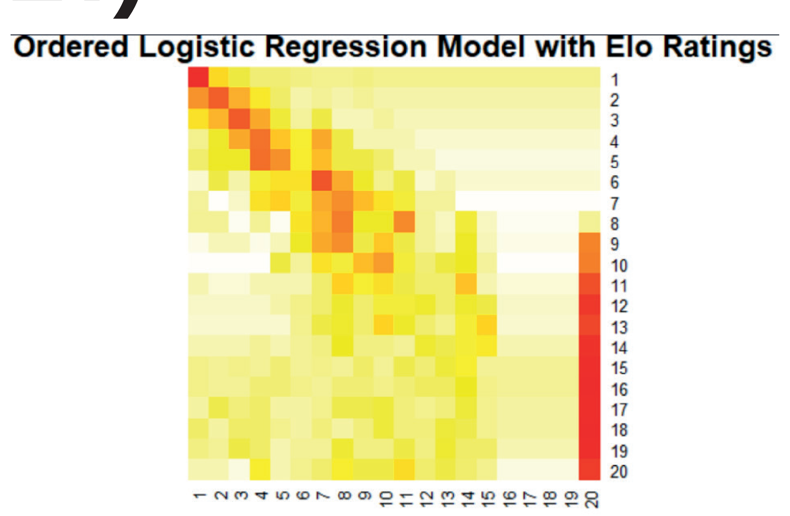
Results (2017-2021)



R-Squared: 0.636



R-Squared: 0.322

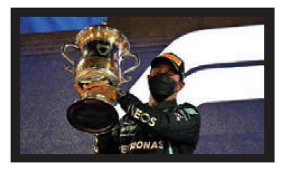


R-Squared: 0.554

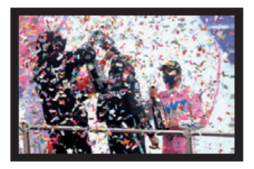
- Overall, the best model was the Ordered Logistic Regression. It predicts the top 10 drivers quite well, and has less confusion near the rear of the field.
- The Elo only model does not do a very good job. Though there is a vauge diagonal, it is much less focused. The R-squared also suffers.
- The combination model is surpisingly worse than the control model. The running average method seems to be better than Elo ratings in this case.

DNF Issues

- The big red lines on the ends of the confusion matrix heatmaps are mostly caused by DNFs (did not finish) results.
- The model is not able to predict for people not finishing. There are a variety of reasons this could happen; a crash, a mechical failure, or a disqualification due to misconduct.
- Removing DNFs increased variability explained by up to 20%, but this is not really a fair thing to do as DNFs are just a part of racing.
- Future research could look into creating some sort of likelihood predictions of certain drivers crashing or certain cars breaking to try to better account for this.



Conclusions



- The combination model and Elo in general performed worse than expected. It was hypothesized that combining data from each race weekend to approximate car perfomance at specific tracks with an Elo rating to try to approximate overall driver performance would make a better model. This was not the case.
- Racing is inherently random. A rigid model can never do a great job in this kind of situation, but this one seemed to do about as well as can be expected.
- Future projects should look into improving the implementation of Elo for multi-team instances like Formula One, and into trying to quantify the frequency of certain car/driver DNF results.

References

Jenkins, Mark. "Technological Discontinuities and Competitive Advantage: A Historical Perspective on Formula 1 Motor Racing 1950-2006." Journal of Management Studies, vol. 47, no. 5, 2010, pp. 884–910., https://doi.org/10.1111/j.1467-6486.2010.00928.x. Mittal, Raghav. "What Is an Elo Rating?" Medium, Purple Theory, 6 Nov. 2020, https://medium.com/purple-theory/what-is-elo-rating-c4eb7a9061e0.

Pelánek, Radek. "Applications of the Elo Rating System in Adaptive Educational Systems." Computers & Education, vol. 98, 2016, pp. 169–179., https://doi.org/10.1016/j.compedu.2016.03.017.