**Project 3 DAM football forecasting- plan & draft**

**Title - Football forecasting - Can the bookies be beaten ?**

**Abstract:**

This is done most easily once we have finished the rest as it is, after all, a brief summary!

**Introduction:**

Betting syndicates, such as Bet365, make billions of pounds a year through offering odds that will entice customers to place bets, but put the bookies in pole position to win. Central to this is a forecasting model; if one can accurately predict future football scores (via a probability of home win, say) then they may offer odds which are slightly above this in order to tip the balance into their favour. This paper looks at building such a model, from a rather simple initial template, before building up to a more complex model. The data available are the statistics from all premier league games from the beginning of the 93/94 season until 29/01/2020. Early season data just contains statistics win, draw and loss, and later seasons expand out greatly with more in-game statistics, such as corners, fouls and cards, as well as betting odds for the game in question. Performance is tested via application to real life games, and predicting real life games, as well as using PLL. A good model ought not to be league-specific, and so the model is also tested on other leagues for its accuracy. Such tests include whether the score can be accurately predicted, and considering what bets one can make when they feel, according to the model, that bookies are giving odds advantageous to the customer.

This paper also undertakes consideration of a Poisson process in order to model goals being scored; predicting goals scored for home and away team is equivalent to predicting the score-line. An adaption of this, known as the Dixon-Coles model, is contemplated too. Use of the Elo rating system is also considered when looking at upcoming fixtures.

* **Home Advantage inclusion?**
* **Converting bookmakers’ odds to probabilities?**

**Model evaluation/ performance measures**

It is vital to ensure that our model predicts well, and so we discuss ways to test this (We shall primarily test our models against the bookmaker’s). One such measure is the mean square error of our probabilities to the bookmaker’s implied probabilities. If this is small then this is an indication of success. However, we would also like a way to test our models without the assumption that the bookmakers are doing a good job. As such we introduce the PLL measure, which does not rely on the bookmaker’s, defined via:

This result is always non-positive since log is non-positive for x in [0,1]. As such, we wish to maximise this via making it least negative as possible. A PLL near 0 indicates the model has predicted the event with high probability, when the outcome of that game is realised. This still allows for comparison to bookmaker’s however, since we can merely compare PLL ratings. For data of varied lengths one can take mean PLL as a measure instead. Since our dataset concludes such betting odds, we can immediately realise the following performances of respective betting companies:

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Company | Bet365 | Betway | Stan James | Interwetten | Ladbrokes | Sportingbet |
| Mean PLL | -0.9554 | -0.9574 | -0.9611 | -0.9585 | -0.9630 | -0.9644 |

We note that the results are very similar between the bookmakers-perhaps they use largely similar models to make their predictions! From this we can see a model with mean PLL of -1 would be of similar quality to the bookmakers’, and models that approach values greater than -0.95 over large testing datasets will actually predict better than bookmakers.

**Methods:**

**1.Generalised Linear Model**

In addition to the aforementioned home advantage, one may wonder what other factors may influence the probability of each result? An obvious one would be the teams involved. But with this comes complications: Manchester City of late have been one of the most dominant teams in Premier League history but were mediocre at best before their financial takeover in 2008. So we may have an interaction between the team playing and the time period that the game takes place. In addition, there may be a situation where Team A and Team B have met 20 times previously, and Team A have won 19 and drew 1. In this case, there is possibly an interaction between the two teams. A more serious flaw is that teams that enter the league for the first time ever (recent examples are Brighton and Bournemouth) have such little data available that, initially, it is hard to gauge how they will perform. A useful way to avoid some of these issues is to use Elo rating.

**Elo Rating**

Elo rating is a metric used to quantify the relative skill of players or teams in games, which leads to a power ranking of the competitors. Originally applied to games like chess, there is research into applying such theory to football. The inclusion of Elo ratings to this report are a result of a few useful qualities:

1. A model which depends only on Elo ratings and not historical data from specific teams’ results should, in theory, be easily generalised to other leagues (one of the key ideas of this report). All that is necessary is one season’s data for training and continuous recordings from this point (does not matter which season or league). For example, if we want to use such model to predict the 2019/20 Bundesliga season, we require full and continuous Bundesliga data from at least 2018/2019- gaps in data is not consistent with the discrete differential equation involved in the Elo rating method. We must also update Elo ratings throughout the predicted season so that every prediction has up-to-date comparisons.

2. A consideration mentioned at the beginning of this section was that certain teams (like Manchester City) have performed very differently across the years. Since Elo ratings alone do not take into account the team in question, only the difference in recent skill between two teams, the issue of historical data being used for current games is alleviated.

Now that some justification behind the use of Elo rating in this report has been given, we will discuss how it is calculated. First of all, there are many adaptations of Elo theory but, in our case, we will consider a system implemented by Stuart Lacy (research fellow at the University of York) that always has a 1500 mean rating. This is a result of the system being zero-sum. That is, if a winning team gains a certain magnitude of rating, the losing team loses the same magnitude of rating. We are concerned with the difference in Elo rating between the home team and away team:

A common assumption is that a team’s performance is a random variable centred on its rating. The underlying distribution of the performance is often taken to be the logistic distribution. The reason behind this is that upsets in football are fairly prevalent and the logistic distribution has larger tails than the normal (for example). Since both Elo ratings in the equation above are logistic about the same mean, is also logistic but centred about zero. One might therefore assume that a higher value ofwould indicate a higher chance that the home team wins. In fact, this is the case. Using a logistic curve, an expected outcome is calculated by the CDF:

Clearly, would indicate a home win (whenis large and positive), would indicate a draw (close to zero) and would indicate an away win (is large and negative). One thing to note is the parameter 400. This scaling constant has previously been found to work well on football Elo ratings and means that .

But we already know that a home advantage exists. To this end, should we really expect two teams of equal skill to draw? One could certainly claim that in this event, the home team should be more likely to win. In fact, Lacy adapts the model above by including a home advantage, labelled HA:

HA is calculated by looking at historical data. Assume that both teams are of equal quality. Then, so:

Using our data, we find that . Then (this is a slight deviation from the value Lacy used). The update of Elo rating is governed by the equation:

Where K and G are variables to be discussed, the dash represents the updated Elo rating and is the observed outcome: if the home team wins, if teams draw and if the away team wins. Consequently, if a team ‘punches above their weight’ and is much larger than E, their Elo rating will increase more than if they beat a team far inferior.

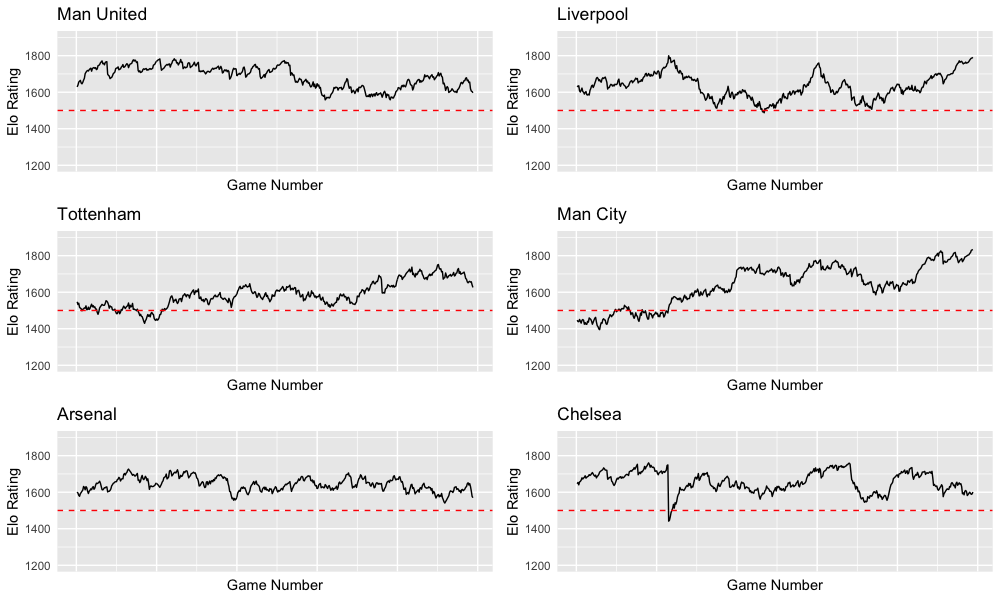
The simpler parameter, K, is arbitrarily chosen in most models as 20. The effect of this parameter is to scale the importance of each result. For tournaments with less games, K may be large but for European leagues, teams play many games so K can be relatively small.

The more difficult parameter is G, which is dependent on the margin of victory (MOV). We shall use the adaptation that Lacy uses in his article. Sparing most of the details, Lacy’s formula for G contains a few key ideas. G will be affected more when the MOV goes from 1 to 2 than it will be from 7 to 8 (for example). That is, if G were continuous, ; the increase of G becomes less as MOV increases. This decreases the importance of huge victories- winning by 8 goals is hardly more convincing than winning by 7. In addition, there will be a penalty added to G when is large. When one team is far better than another, the effect of large MOV will be less than if the teams are similar in ability. The value of G will depend on MOV like so:

There is one last issue with this model. Unlike other sports (like basketball), every year three teams with the lowest points will be relegated from the league and three teams from the Championship will be promoted (2 automatically and 1 through play-off games). So the teams in the league change annually and our model must find a way to deal with this. Since the teams relegated from the Premier League and the teams promoted from the Championship bridge the gap between the two leagues, the most informed assumption would be that they are of similar quality. Of course this assumption can be false in some cases, but it is the best simple assumption to make and their Elo rating will be corrected within a matter of games regardless. Justification for this is that there is only one occasion since 2006 where no promoted teams were relegated in the next season. Most of the time, at least one of the teams promoted face immediate relegation in the following season. To ensure that the mean of 1500 rating is consistent throughout, we can assign the average Elo of the three relegated teams to the three promoted teams at the start of the following season.

In addition, due to factors like end of season form and transfers, the gap in ratings will be pulled towards the mean using . Hence, the hierarchy of teams is maintained but the gap in quality is reduced to account for the uncertainty at the start of the season. At the end of each season, the ratings will be pulled towards the mean and subsequently the promotion/relegation theory previously mentioned will be applied.

Having applied this measure to our Premier League data (after first training it on the 2005/06 season), we had a dataset containing results from 2006-2019, along with the Elo ratings of the home and away teams involved. To give this theory some colour, an example of Elo over time can be visualised by comparing the ‘Big 6’ over time:



The dotted red line represents the league average Elo. It is encouraging to see that Manchester City and Liverpool have the best Elo ratings in recent years which definitely reflects reality.

**Building the model**

With three outcomes, a home win, a draw and an away win, we require a multinomial output with away win acting as a reference level. The generalised linear model with home and away Elo as the only covariates gives equations of the form:

Using data from the 2006-2019 seasons, we have 4940 games. Luckily, our model doesn’t rely on the time that data is recorded or the teams involved. Our model is only reliant on Elo ratings before a game instance occurs. We can therefore split training and testing data in any manner we like. For ease, we shall take the first 70% of data (3500 games) to train and the remaining 30% to test. This gives estimations to the parameters of our predictions as:

We have three equations with three unknowns, which is solvable provided we know the Elo ratings of each team.

Despite exploring other variables to include in this model, there are several reasons for leaving this one untouched. The first idea was a variable for each team. I have no doubt that the inclusion of a team covariate would improve the PLL for certain games in a season. However, we arrive at the same conclusion as we did at the beginning of this section: a model that depends on a team variable across many seasons will only be good for some teams. Many teams have limited data points available. Some have no data at all- if we wanted to predict the 2017-18 season from any of the previous years’ data, one couldn’t possibly calculate a parameter for Brighton, who entered the league for the Premier League for the first time that season. Moreover, even if we could calculate a parameter for each team, it would have to be weighted so that recent games have more of an effect. And even if we could implement such a weighting system, how much more of the variability of the response variable would be explained by a team variable, that Elo hasn’t explained already?

In an attempt to create more abstract variables, one could try and create variables that represent the ‘form’ of each team at each game. If we define form as the accumulation of each game metrics (goals, goals against, shots, shots against, red cards etc) across the last five games for each team, would this be useful in our model? In short, none of the form metrics are classed as significant on the generalised linear model, once the Elo ratings have been accounted for. Regardless, for each form metric to be calculated, we must wait around fifty games in a season for each team to have played five games each, in the first place. So even if it had an effect on the model, it reduces our dataset and means our model won’t make predictions at the start of the season which is something we’d want a general model to do.

In order to keep the model general and versatile, we leave only the Elo ratings to predict games. This allows us to predict every game of each season for many different leagues. Later parts of the report will look into more specific predictions.

**Elo Model Evaluation**

Applying this model to the testing dataset gives a mean PLL value of -0.9552. Clearly this model performs similarly to the bookmakers’, at least under the PLL measure. Using the same training data, Bet365 (who historically perform the best under the PLL measurement), achieved a mean PLL of around -0.941.

This model approaches a similar degree of accuracy in its predictions as the betting companies do. This is a very positive result- this model depends only on Elo rating and nothing else, whereas huge companies like Bet365 with many employees will be able to adapt and tailor their odds to each and every game. They will undoubtedly have more data available, as well as information like transfers and appointments of new managers which can have an immediate effect on the outcome of games, something Elo rating doesn’t necessarily react as quickly to.

So, whilst this model hasn’t beaten the bookmakers in our predictions, there are a few positive results:

1. It performs significantly better than simple models; it has a mean PLL 11.2% better than the home advantage alone.
2. It isn’t considerably worse than the best of the bookmakers with a mean PLL rating of only 1.5% lower.
3. This model is highly generalisable to other leagues, one of the aims of the report. It is fairly low maintenance, requiring only the teams involved and the final score as data.

To elaborate on point 3 further, we tested this theory on the 2018/19 Serie A season and achieved a mean PLL of -0.9745 compared to Bet365’s -0.9448. To an extent, our hypothesis was true: our model still does a good job of predicting (around 3% worse PLL than the bookmakers) compared to simple models like the home advantage alone. However, it must be noted that it performs worse compared to bookmakers on the Serie A than on the Premier League. Perhaps this is because we trained the model on Premier League fixtures and assumed the home advantage is the same in both leagues. Furthermore, parameters in the model may have been influenced by the ‘predictability’ of the league it was tested on. If one league is more predictable - teams of higher quality beat teams of lower quality more frequently - perhaps the Elo rating would influence the predictions more.

If one wants to have a better insight into the probability of each result happening, the Elo model is useful. It provides far better insight than blindly expecting the home team to win. In this regard, and because no football predictions can be made with 100% accuracy, it does a good job of answering the question posed by the report.

The model is extremely flexible. It can be used to predict any game of any season pretty well. However, it must be noted that it performs marginally better on the Premier League because this is the data the model was trained on. So there are some limits on the use of the model (though it was never claimed that this model was highly predictive in the first place). One could train the data on the league one wishes to predict, but this removes the intrinsic convenience of using such model.

To conclude this section, the use of the Elo GLM is far more informative than simple models and extremely versatile in its use. However, it fails to reach quite the same quality in its predictions as the bookmakers do, especially when predicting results of other leagues. The idea of this model is not to be finely tuned to each and every game, but to be useful in fulfilling one of the aims of this report: to predict any football match, regardless of league. But what if one requires the kind of accuracy in predictions to beat the bookies? Betting syndicates surely wouldn’t place all of their money into a general model like this one. In order to make money from the fine margins in the betting markets, our model must be highly specific to each game. We must use models that aren’t necessarily as easy to apply but are able to exploit weaknesses in bookmakers’ predictions. In the next section, we will explore models that are of this nature.

**2.Poisson Regression and its improvement**

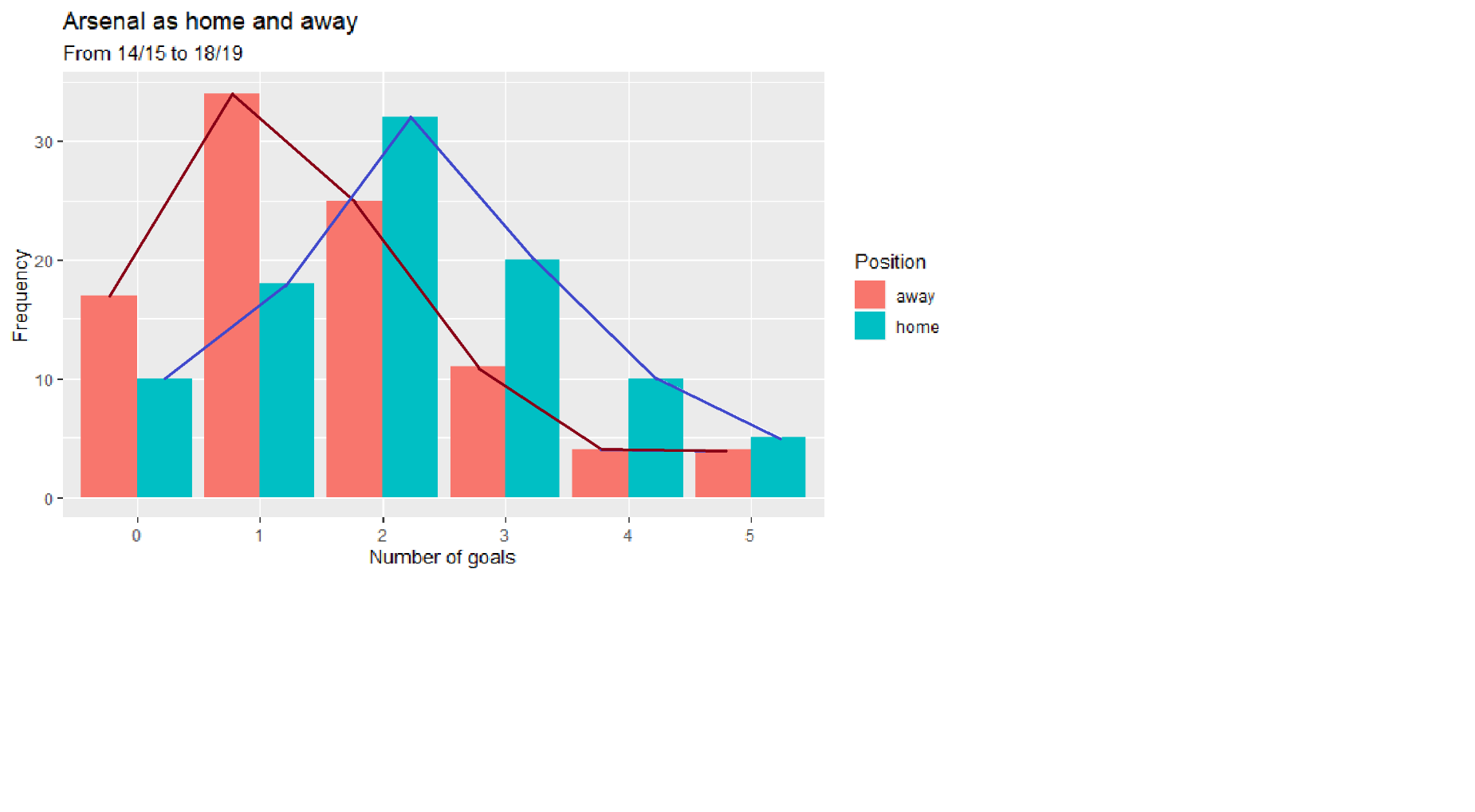
In general, the idea of the Poisson regression model here is to use the maximum likelihood to formulate parameters to characterise each teams’ level of offensive and defensive skill and hence to predict the probabilities of each game result.

**Poisson Distribution**

Before building the model of Poisson regression, it’s necessary to assume that the number of goals for each team follow independent Poisson distributions. In addition to the plethora of academia using such distribution, this assumption is further justified by the following idea: in theory, one of the key concepts in football is the possession of the ball. During each possession, a team will try to attack and score against the defence of the opposition, with a small probability of success (scoring) of . Since the number of possessions for each team is ideally very large and, under the assumption that is constant and each attack is independent, the number of goals follow the binominal distribution which approximates the Poisson distribution according to Poisson limit theorem.

Poisson limit theorem:

If this isn’t enough, the Poisson distribution can also be shown by the data. Here is a plot of the number of goals Arsenal score, broken down into home and away games. Clearly, the goals generally follow the Poisson distribution:



**Poisson regression model**

Before showing the model, the following assumptions are made:

1. The number of goals for each team follow a Poisson distribution

2. The number of goals between each team and their opponent is independent

3. Each score-line is independent from match to match

Here is the model:

(1)

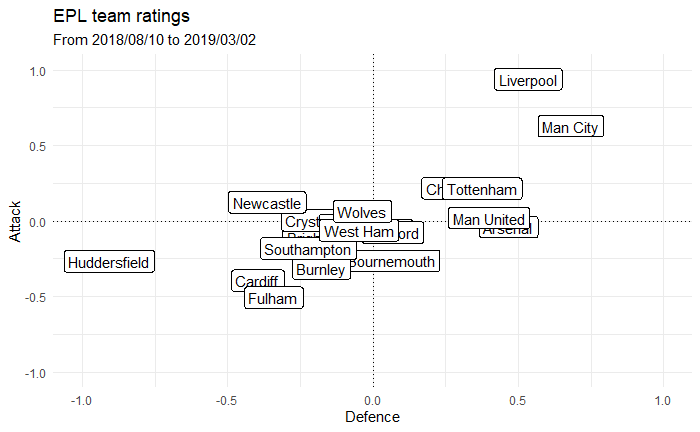
and are the random variables for the number of goals scored in the game where team plays against team.

From the assumptions, these two random variables are independent so that the joint probability of the home team scoring goals and the away team scoring goals is the product of the two independent probabilities.

There are several ways to evaluate the expected value of parameters and . One way is to express it by the attacking ability of the home team, the defensive ability of the away team and the existence of the home advantage, which has been proved that it affects the result of the game.

Here it is:

Where , and represents the attack and defensive ability of the team respectively and corresponds to the effect of the home advantage. To visualise the values of and for each team, here is a plot for the 2018/19 season:



After substituting the expected values into the pmf(1) and take the Log, the expression turns into:

If is a game instance and games are observed in total, then the log-likelihood of these k games would be:

To avoid multiple combinations of parameters which may produce the same model, this constraint is added:

After using the optimisation method to calculate the maximum log-likelihood, the value of each of the parameters could be obtained. The probability of each result (home win/draw/away win) can be forecasted like so:

**Dixon-Coles Model**

Different from the assumption in the Poisson regression model, the Dixon-Coles model assumes that the score-lines 0-0, 1-0, 0-1 and 1-1 are not independent. Therefore, one probability modified function is to be added to the Poisson regression model.

and the Dixon-Coles model is:

The modified function is the only difference between the simple Poisson model and the Dixon-Coles Model, which means the ways to calculate the parameters are generally the same (by MLE). What is the meaning of ? Does it depend on the data? Is it constant for every game?

**Time weighting**

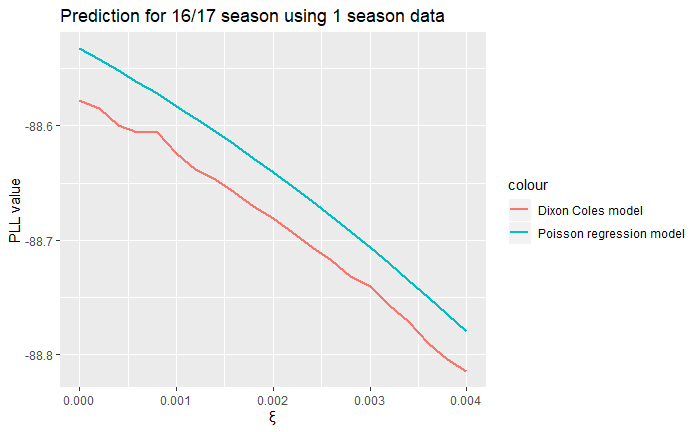
During building the model for game forecasting, the more recent results are more significant as they better reflect the current level of the team. Therefore, the weight function is introduced and used here:

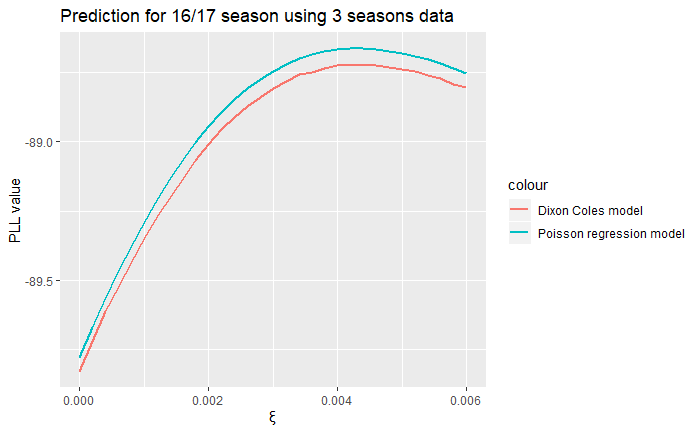
Where the is a constant and is the time to predict. The meaning of is the length of time previous to current day for which one deems a game to have no effect on the result of a current fixture. Games that approach this time will have less and less effect on our model. The former log-likelihood for Poisson regression turns into:

Therefore, for each change in the value of , the value of the parameters in the model will be different by calculating the MLE. Following this, a different mean PLL for each model can be calculated. The  with the highest mean PLL may represent that such model could predict the results best and that should be used for the foregoing forecasting.

In addition, it could be seen that the weight function depends on . However, is just a point, not an interval, which means the weight function needs to be shifted after each prediction when forecasting more than one date. To select the correct for the time weighting function, after a single game instance is predicted on a certain day, the current time is updated, a single game at the back end of our data is removed from the dataset and the single game we have just predicted enters the dataset in order to predict games for the next fixtures. This maintains the size of the dataset after every prediction and flushes out games that now exceed and hence loses its validity.

Here are plots of the mean PLL for predicting the last 100 games in 16/17 season using the different given by the Poisson regression model and the Dixon-Coles model. The first plot is using 1 season’s data and the second uses 3.





It can be seen that, in order to maximise the mean PLL, both Dixon Coles model and Poisson regression model could predict best without adding time weight when using 1 season’s data to build the model. However, when using the 3 seasons data to build the model, the of weight function should be set as 0.0044 for the best predictions.

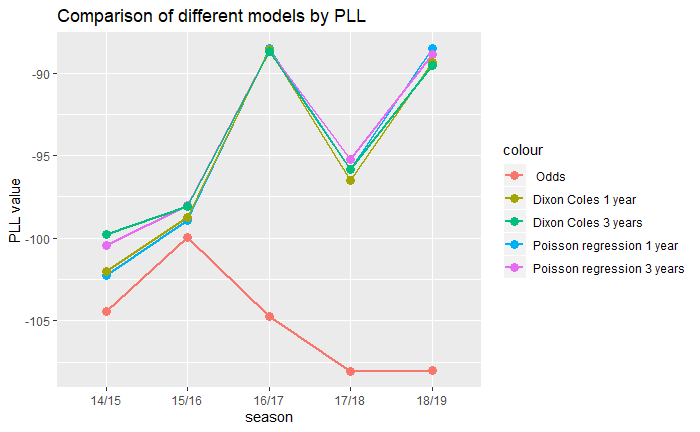
**Predicting the Results**

To predict the last 100 games in 5 seasons, Dixon Coles model and Poisson regression model using 1 season and 3 seasons data are built and mean PLL is calculated to evaluate the different models. In addition, the PLL of the implied probability given by bookmakers’ odds are also calculated as a reference.

To justify predicting only the last 100 games of each season, it can be observed in football that there is an underlying uncertainty of results at the start of the season due to obvious changes in personnel at certain clubs over the transfer window (new signings and managers), newly promoted teams of unknown quality and perhaps, more conjecturally, a ‘new season mentality’; results at the beginning of the season can be erratic, with bad teams of the previous season starting with the optimism of a clean slate and good teams potentially losing their end of season form.

In this section of the report, we are concerned only with predicting games with the highest possible accuracy. Perhaps for the previous part of the season, we would use the Elo generalised linear model since this, to some capacity, combats the problem associated with such uncertainty with the help of a few assumptions, with the drawback of some accuracy.

Here is the comparison plot of each model:



From the plot, it could be seen that all the models are able to predict much better than the odds of the bookmaker and the predicting ability between each model is very similar. There was not a model that could always obviously predict better that the others. Furthermore, the Poisson wqqwble below shows the mean PLL of these 4 kinds of models, which would be helpful to compare with the models generated by the methods of different types.

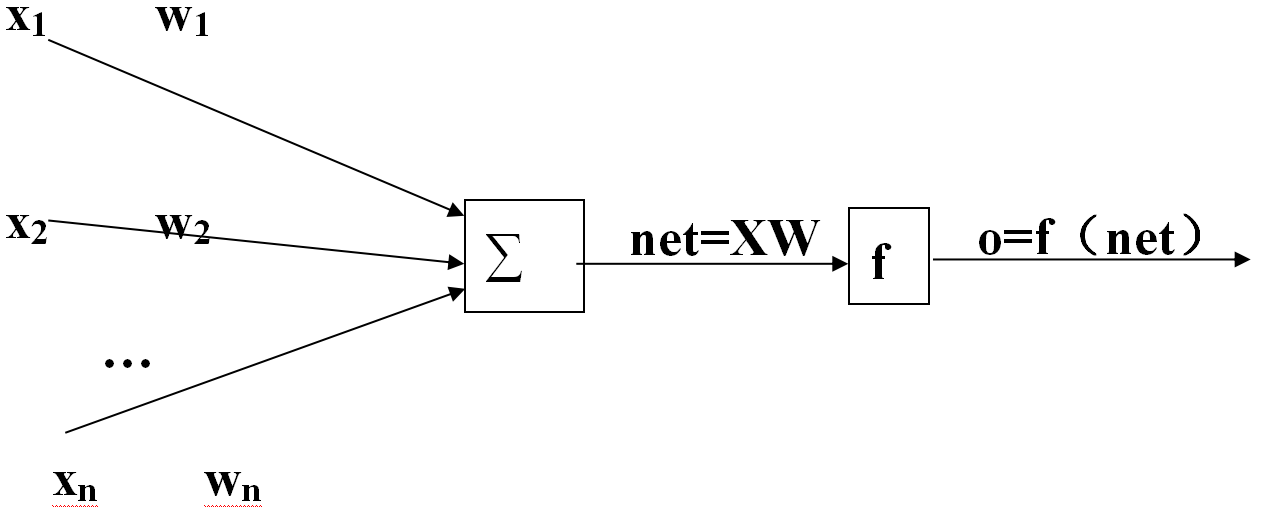
|  |  |  |
| --- | --- | --- |
|  | Dixon Coles model | Poisson regression model |
| 1 year data | -0.9504066 | -0.9481608 |
| 3 years data | -0.9439942 | -0.9424600 |

**3.BP ANN**

Artificial neural netwo**rks (ANN) are a simulation of human neural networks. More directly, it is a mathematical model, which can be achieved by computer or electric circuit. It has been found to be an effective way to research Artificial intelligence(AI). Simpson (1987) claimed ANN is** a non-linear directed graph. The graph contains some sides with weight which can be changed, and the ANN is able to find the pattern from incomplete or unknown input. Artificial neurons are the basic unit of ANN, which is a simulation of biological neuron. It has six characteristics:

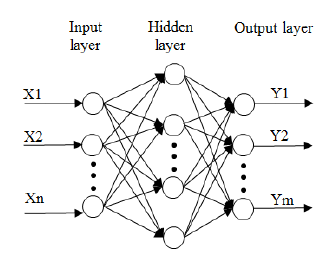
1. Connect the neurons
2. The strength of connection between neurons determines the strength of signal transmission
3. The connection strength between neurons can be changed with training
4. Signals can be stimulating or inhibiting
5. The cumulative effect of signals received by a neuron determines its state
6. Each neuron can have a threshold

If we have inputs, and weights, , the net input is given by and passed through an activation function . This process can be visualised below:



Werbos (1974) raised the back-propagation learning algorithm for multilayer feedforward networks which was developed by Rumelhart in 1986 (Rumelhart et al. [1986](https://link-springer-com.ezproxy.nottingham.ac.uk/article/10.1007/s00477-007-0181-7?shared-article-renderer#ref-CR16)).

Back-propagation artificial neural network (BP ANN) often contains one input layer, one or more hidden layers and one output layer. There are a certain quantity of neurons in every layer. And the neurons are connected by certain weight to next layer’s neurons. Once these neurons work together, the network can solve problems. Each neuron accumulates the signals from previous layer, and deals with the signal by an activation function (often a sigmoid function).



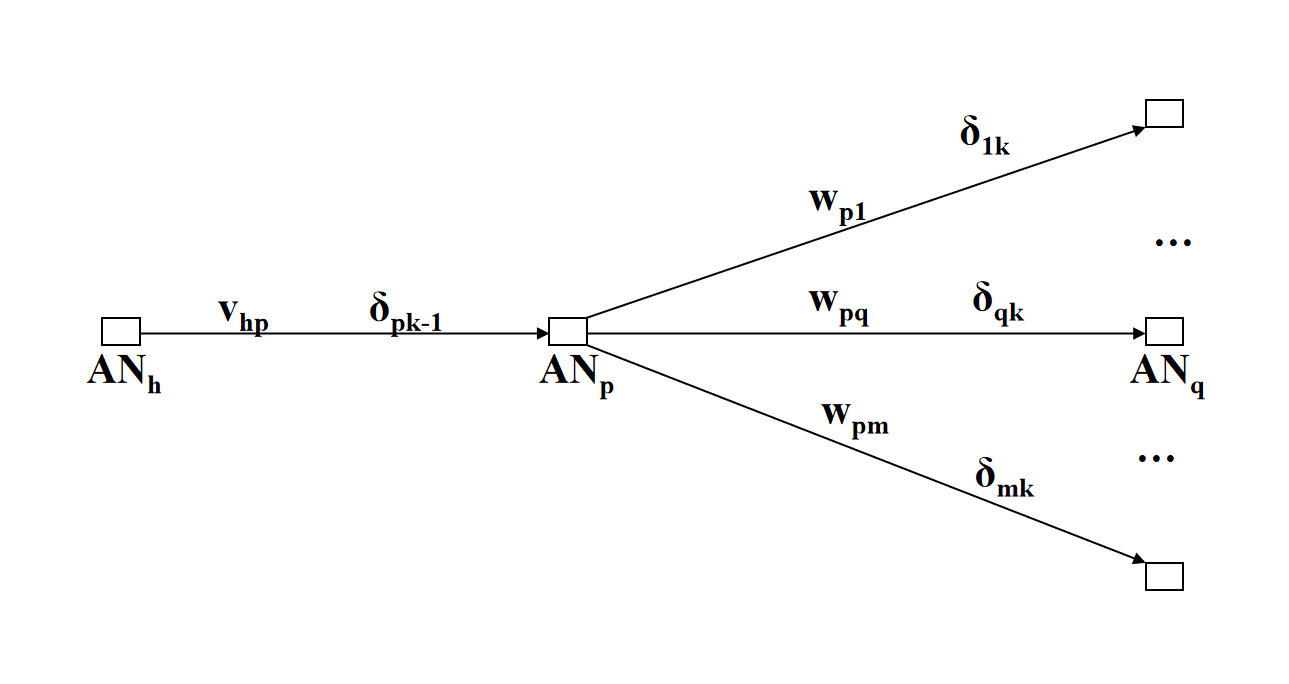
Basic characteristics:

1. Network contains layers
2. Existence of Connection Matrix:
3. Sample set
4. Neurons in layer k is

According to the samples in the sample set, calculate the actual output and calculate the error between the predicted and observed output, then adjust the weights , and repeat this procedure until the where .

The adjustment of connecting weights of (what does this mean?) can be found by using the steepest descent method (Rumelhart et al. 1986) as below:

Where is learning ratio, which can control the change of step. In addition, the adjustment of output and hidden layers is given by:



After searching the effect of every covariate, we use the data from Premier League in 17/18 (380 matches) as the training data. If one finds the cumulative statistics of the first 280 matches in the season, one finds that the significant inputs that affect the results of the game are the average of [Home shots (HST), Home corners (HC), Home red cards (HR), Away shots (AST), Away corners (AC), Away red cards (AR)]. For example, if Arsenal are the home team, calculate the average HST, HC, HR, If Liverpool are the away team, calculate the average AST,AC,AR. More clearly:

* Input:[Arsenal[HST],Arsenal[HC],Arsenal[HR],Liverpool[AST],Liverpool[AC],Liverpool[AR]]
* Output: if home win: [1,0,0], if draw: [0,1,0], if away win: [0,0,1].

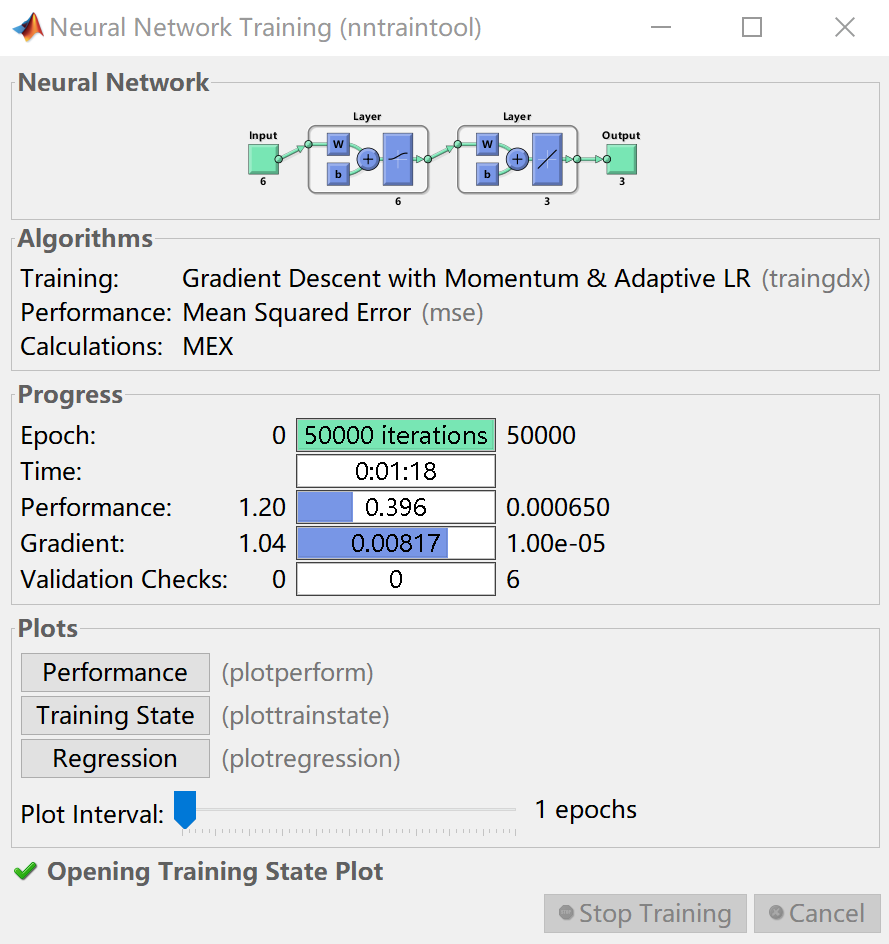
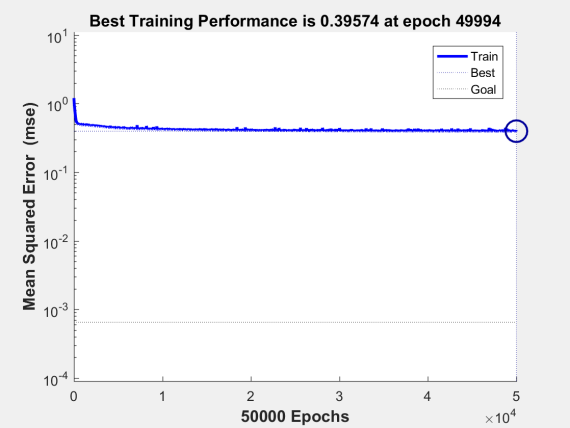
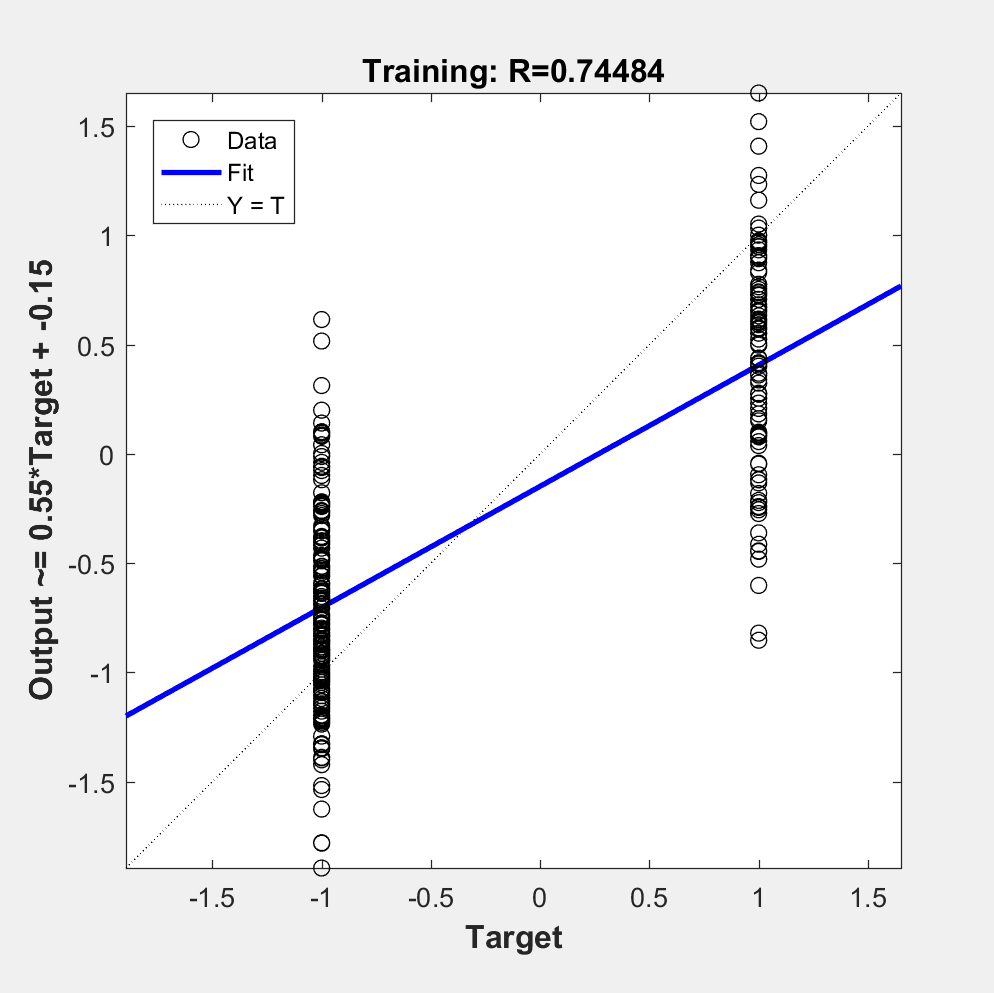
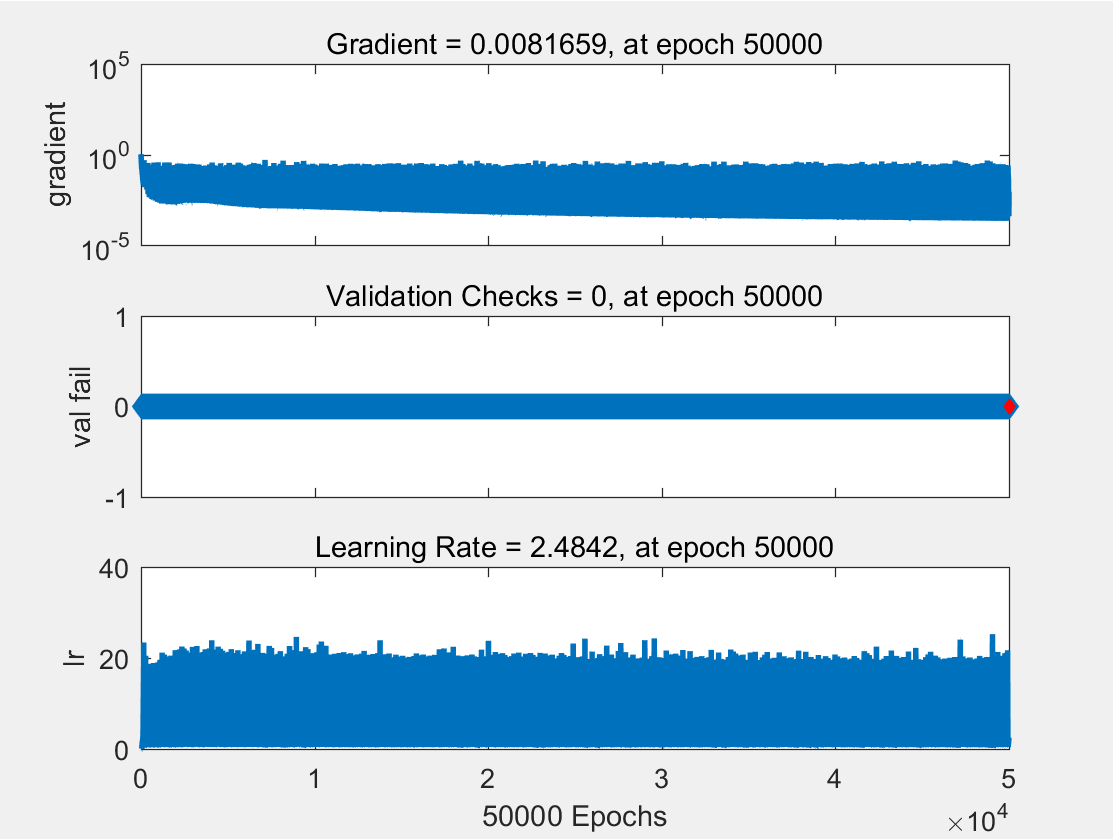
For testing, this model selects the data of premier league in 18/19 (380 matches). We use the first 280 matches to train the model, and the last 100 matches test (like the previous Poisson model). In our network, the number of neurons in hidden layer and output layer as 6 and 3 respectively connected by log-sigmoid function and ‘purelin’ function, and using ‘traingdx’ as the training function for gradient descent. Using Matlab, some parameters used are shown below:

trainParam.show=2000

trainParam.Lr=0.01

trainParam.epochs=50000;

trainParam.goal=0.65\*10^(-3)



Explain what the graphs show?

When the output is found, employ the softmax function to transform it to a probability of home win, a draw or an away win. The standard (unit) softmax function that takes : is defined by the formula:

for and

After analysing the result, we find that the BP ANN predicts 57% of the last 100 matches of the 18/19 season and has a mean PLL=101.2. as the output showed in appendix.

**Results:**

-

**Discussion:**

Limitations

-Teams in transition:

Whilst the data used in the formation of the model are extensive they are not exhaustive; there exist other factors which may affect the outcome of a game, which may not be simple to incorporate into a model. An example of this (with varied severity) is player suspensions and injuries;many teams are highly dependent on a couple of key players, whose absence could negatively affect a team’s chances. For instance, a striker who is a team’s main goal contributor will likely be missed. This is not particularly simple to consider when modelling, but may be considered by betting syndicates, and thus may explain why bookies appear to do better in some cases.

Similarly, it is often the case that a team playing particularly badly will look to change their manager. Whilst our model will not anticipate or identify this, “new manager bounce” is very often seen at a club, whereby a turnaround in form occurs when a new manager takes over. This may be another factor that is neglected, but whose consideration ought to be able to improve our model. These game-specific factors may sometimes allow the bookies an edge.

One should also consider the data available. If this approach were used on a team who were newly promoted, there would be little-to-no data to make use of yet, and similarly with newly relegated teams. This drawback is not particularly easy to negotiate, but one that would only serve as an issue for a very early period of a season, until enough data are gathered.

**References:**

1. https://stuartlacy.co.uk/2017/08/31/implementing-an-elo-rating-system-for-european-football/
2. Werbos P (1974) Beyond regression: new tools for prediction and analysis in the behavioural sciences. PhD Thesis, Department of Applied Mathematics, Harvard University, Cambridge
3. Rumelhart DE, Hinton GE, Williams RJ (1986) Learning internal representations by error propagation. In: Rumelhart DE, McClelland JL (eds) Parallel distributed processing: exploration in the microstructure of cognition. MIT Press, Cambridge, pp 318–362

**Next Week:**

* Write up some of the ANN section in this google drive, including the selection of variables (sent in Whatsapp). We will check the grammar, like asked.
* Write up the home advantage somewhere early on