The Numbers Behind the Numbers

Gil Weissman

September 2020

Table of Contents

Part 1: Introduction

As a typical Israeli teenager, football has always been my favorite sport, both for watching and playing. I have watched football matches since an early age. The local league at the beginning, followed by much more interest in the main European leagues. Two years ago, I started the Data-Science BSc in the Technion. Watching some really inspirational movies, like Moneyball, and start working with big data, immediately made me think of a possible connection between those two big topics in my life.

Another thing that made me think of a way to combine Data science algorithms with football, was a video I watched, where a google programmer explains how they correctly predicted 14/16 games in 2014 FIFA World cup by using machine learning algorithms and football datasets. It gave me an actual proof that this combination can work, and gave me a first idea of how to do it. I highly recommend watching this video on YouTube - Google Predict football matches

The biggest online sport betting website today is bet365. It has betting possibilities in almost every sport you can think of. Not only that it gives you the opportunity to bet on almost every statistic in the game – for example in football you can bet on the winning team (or draw), correct score, match goals, match corners and many more, but also allows you to place those bets during the matches, according to their progress. Unfortunately, as an Israeli resident I am not allowed to bet there. In Israel, Winner has a monopoly on the sport betting field, and therefore do not have to fight hard to satisfy its captive customers, which results in low winning rates, not many betting possibilities and an old-fashioned website with a pretty bad user experience.

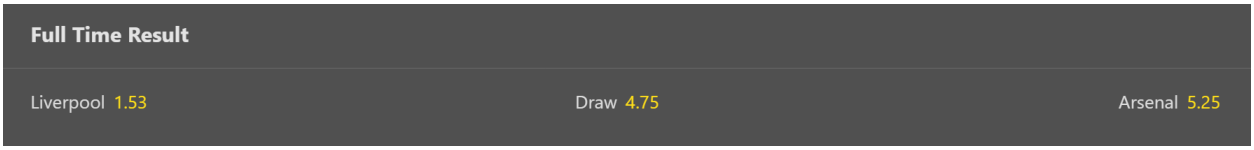
My main goal in this project was to find an algorithm that can actually make money by placing bets on football matches. And indeed, I successfully programmed some classifiers that correctly match games to game results (Home win, Draw and Away win) which makes money over time by placing bets.

On the next pages, I'm happy to share with you the process I've been through on my way to find those profitable algorithms.

Part 2: The 3 - Way Betting Line

The most common football betting line is a 3-way line (Home team, Draw and Away team). When betting on the 3-way money line you have three betting options. Your selected option must be correct for your bet to win. For example, if you selected the Away team and the event ends in a Draw, your selection on the Away team loses.

There are many ways to analyze the 3-way betting line, I will present some of them using this example from [bet365]:



So, what do we see in this betting line?

Basically, it means that by betting X amount of money on Liverpool, if Liverpool wins, you get your X back plus $0.53X$, but if they do not win you lose the X you bet. Similarly, by betting Draw, if the final result is a Draw you get your X back plus $3.75X$, but if the game ends up with a win for either the Home team or the Away team, you lose the X you bet. And as you probably already guessed, the same goes with betting on Arsenal. You can also notice that Liverpool is written on the left side of the bet, which indicates that they are the Home team on this match.

Now, let's explore the numbers written in the bet line, because they can tell us a lot about the game. The numbers in fact, represent the probabilities for the different possible events to happened. The Implied probabilities are very easy to calculate by the simple function: $Imp.prob = \frac{1}{bet\ ratio}$. Which means that the probability of a Liverpool win is 65.3%, the probability of a Draw is 21% and the probability given to an Arsenal win is 19%. But what does it mean? People have a problem of understanding the meaning of probabilities, it actually has a deeper meaning rather than Liverpool is the favorite team on this match. You can find a way of understanding those probabilities on the next pages, but at the mean time keep in mind that summing up those three implied probabilities results in more than 100%.

Another way of looking on this 3-way betting line is as a zero-sum game (not exactly…) between the better and the casino. A zero-sum game in economics and game theory is a mathematical representation of a situation in which each participant's gain or loss of utility is exactly balanced by the losses or gains of the utility of the other participants in the game. If the total gains of the participants are added up and the total losses are subtracted, they will sum to zero. A zero-sum game with two players is a situation where the loss of the first player is the payment to the second one, and vice versa. In such a game each player has a finite set of actions he can play, and the payment is decided according to the actions of both the players. You can read more about zero-sum games here - zero-sum game.

Well, its pretty easy to guess that the set of actions the better has is: placing a bet on either Home win, Draw or Away win (or not placing a bet at all). But what is the set of actions that the casino has – that's not that obvious… Remember I told you it's not exactly a zero-sum game, well its basically because the casino set of actions is actually the "reality" set of actions – the final result can be either Home win, Draw or Away win. The reason why we would like to think of the bet as a zero-sum game is that zero-sum games have been researched for decades as part of the game theory field, with some really effective solutions that have been found, which we can implement on our problem. The common way of looking at a zero-sum game is as a table:

		Better – the column player		
		Home win	Draw	Away win
Casino – the rows player	Home win	-0.53, 0.53 (block 1)	1, -1 (block 2)	1, -1 (block 3)
	Draw	1, -1 (block 4)	-3.75, 3.75 (block 5)	1, -1 (block 6)
	Away win	1, -1 (block 7)	1, -1 (block 8)	-4.25, 4.25 (block 9)

The table represents the possible actions of the two players, and the final outcomes derived from them. In each of the 9 internal blocks of the table, the left number is the utility of the casino, aka the rows player, and the right number is the utility of the better, aka the column player. For example, block 6 represents

the situation where the better bets on the Away win while the final result is a Draw. On that case, the casino has won the better money, while the better has lost the money he bet to the casino.

For two-player finite zero-sum games, the different game theoretic solution concepts of Nash equilibrium, minimax, and maximin all give the same solution. If the players are allowed to play a mixed strategy, the game always has an equilibrium. An equilibrium in a game is a situation when given that one of the players plays his equilibrium strategy, the best response of the other player is to also play his equilibrium strategy.

Notice that there is no pure equilibrium in the game above, or in other words, there is no block in the 9 internal blocks that satisfies the condition of being an equilibrium in the game. For example, block 1 is not an equilibrium because given that the column player plays 'Home win', the row player would rather play 'Draw' than stay with 'Home win', and block 8 is not an equilibrium because given that the row player plays 'Away win', the column player would rather play 'Away win' than stay with 'Draw'. The same goes with all the other 7 internal blocks. But we said that each zero-sum game like this indeed has an equilibrium. The equilibrium is actually a mixed strategy, which means that the strategy of each player is a combination of his three possible actions.

There is a mathematical way to solve this, but for now we will use this online website that calculates for us the equilibrium strategies - Zero-sum game solver. The website returns $(0.61976, 0.19963, 0.18062)$ for both the column and the row players. The utility for the casino under these strategies is 0.05177, while the utility of the better is of course the negative number: -0.05177

This is not surprise that the better has a negative utility in the equilibrium, since if he had a positive one, he could ensure himself a positive utility by betting on the match, something that the casino will never allow.

Now, assume you have a "magic machine" that gives you the "correct" probabilities for the possible match results. For example, on the $Liverpool\ v\ Arsenal$ game, this magical machine reveals these probabilities: $(Liverpool: 60\%\ , Draw: 20\%\ , Arsenal: 20\%)$. Let's explain first what is the meaning of "correct" probabilities. Given this set of correct probabilities for the

game, does a final result of a Draw make sense? Since the highest probability was given to Liverpool, isn't is supposed to be the final result? Well, the answer is definitely NO! The best way to think about it, is to imagine that instead of the football game, the players decide to play another game, a lottery game. 100 notes are put in a bag. 60 of them are Liverpool notes, 15 are Draw notes and 25 are Arsenal notes. The referee picks randomly one note from the bag which sets the winner of the match. For example, if he picked an Arsenal note, then Arsenal has won the game. To sum up, the meaning of "correct" probabilities for a game are probabilities that represent the chances for each possible result, as the game was not a football game, but a simple lottery game as explained above.

After understanding the meaning of such a "magic machine" that reveals the "correct" probabilities for a match, we can ask whether such a machine can help us maximize our profit from placing bets. Let's take another look at the implied probabilities and "correct" probabilities of the match we are observing. We can see that both for Liverpool and the Draw results, the implied probabilities from the 3-way betting line are higher than the corresponding "correct" probabilities, while for the Arsenal result the implied probability is lower than the "correct" probability. Think about it, if the implied probabilities are higher than the "correct" probabilities, the payment to the winning betters is lower than it should have been, since they are paid as the result had more chance to occur than it really had. On the other side, if the implied probabilities are lower than the "correct" probabilities, the payment to the winning betters is higher than it should have been, since they are paid as the result had less chance to occur than it really had. Therefore, having such a machine is indeed helping, because if we had one, we could look for positive differences between the "correct" probabilities and the implied probabilities, and ensure ourselves a positive utility in the long term (positive Expectation). Understanding the last statements would explain to you why the casino wants the result of summing up the implied probabilities of a bet line to be greater than 100%. Unfortunately, such a machine is not existed, so we have to work hard to find high level Models that will be close enough to this machine. More on this discussion on the Algorithm part.

Another way of using the "correct" probabilities is to find the best response the better can play, given the strategy of the casino, which is {0.6, 0.2, 0.2}. Solving this using a system of linear equations, is quite simple and the result is that the

highest utility the better can earn is -0.018, by playing $(0.634, 0.192, 0.174)$ strategy. The meaning of this better strategy is that if the better wants to bet for example $1000\$$, he bets $634\$$ on Home win, $192\$$ on Draw and $174\$$ on Away win, assuring himself a lost of $18\$$ overall. On this case, even the "correct" probabilities could not assure the better a positive utility from placing bets on the game, though it assures a higher utility than the one we found in the original equilibrium.

The last thing to discuss regarding the 3-way betting line are the changes in the line, before and during the game. The changes during the match are very complicated to understand and depends on many statistics collected during the game. Luckily, we are interested on those occurring before the match starts. There are two main reasons for those changes. The first one could be a big change in the teams' lineups, or a big player surprising absence. The second one is a much more common reason. It's important to understand that the casino wants to win money, regardless of the final result, and its way of assuring this is to manipulate the betters by changing the bet ratios. Explaining this will be much easier using this following 3-way bet line:
$[Home\ win: 5.1, Draw: 3.2, Away\ win: 1.4]$. Assume that people start placing bets on the match, and then the casino notices that a total of $100\$$ bets have been placed on the Home team, while on the Draw and the Away team no bets have been placed at all. On that scenario, if the final result is a Draw or a win of the Away team, the casino earns those $100\$$, but if a big surprise occurs and the Home team wins the match, the casino loses $410\$$. Therefore, the casino wants betters to place bets also on the Draw and the Away team, such that also in a Home win it's utility will be positive. Obviously, higher bet ratios are more worthy for the betters, and lower bet ratios are less worthy. Now, assume that the casino has updated the line to: $[Home\ win: 4.6, Draw: 3.4, Away\ win: 1.7]$. On that case, more people will place bets on the Draw and away win, and less people will place bets on the Home team. A possible result for this situation could be that a total of $100\$$ bets have been placed on the Home team, a total of $150\$$ bets have been placed on Draw and a total of $300\$$ bets have been placed on the Away team. This is an ideal situation for the casino since regardless of the final result, the casino has a positive utility of $40\$$.

Part 3: Data Processing

The most important part in every machine learning process is arranging the data in a way you could use it to train different algorithms and make predictions on new data.

First of all, the initial datasets I used cost me no money, and one can download them for free from this website – [Football Datasets Website].

On this website you can find basic datasets of each one of the primary leagues in Europe and much more. In each of these leagues you can find datasets containing information on each season starting from 1993/1994.

This is an example of such dataset contains data of the Spanish LaLiga from 2019/2020 season (only the first 10 rows and columns):

Div	Date	Time	HomeTeam	AwayTeam	FTHG	FTAG	FTR	HTHG	HTAG
SP1	16/08/2019	20:00	Ath Bilbao	Barcelona	1	0	H	0	0
SP1	17/08/2019	16:00	Celta	Real Madrid	1	3	A	0	1
SP1	17/08/2019	18:00	Valencia	Sociedad	1	1	D	0	0
SP1	17/08/2019	19:00	Mallorca	Eibar	2	1	H	1	0
SP1	17/08/2019	20:00	Leganes	Osasuna	0	1	A	0	0
SP1	17/08/2019	20:00	Villarreal	Granada	4	4	D	1	1
SP1	18/08/2019	16:00	Alaves	Levante	1	0	H	0	0
SP1	18/08/2019	18:00	Espanol	Sevilla	0	2	A	0	1
SP1	18/08/2019	20:00	Betis	Valladolid	1	2	A	0	0

The next stage was taking all of these datasets, each represents one season of one league and merge them into fewer, main datasets, each represents one league, as follows:

Seasons 2009/2010 – 2017/2018 are all merged together into one dataset, which from now on, will be called the Training set.

This is an example of such Training set contains data from 2009/2010 – 2017/2018 seasons of the Spanish LaLiga (only the first 10 rows and columns):

	date	HomeTeam	GoalsFH	GoalsAH	ShootsTargetFH	ShootsTargetAH	B365H	B365D	B365A
1	29/08/2009	Real Madrid	3	2	11	3	1.25	5.5	13
2	29/08/2009	Zaragoza	1	0	8	2	1.91	3.4	4.2
3	30/08/2009	Almeria	0	0	5	1	2.1	3.3	3.5
4	30/08/2009	Ath Bilbao	1	0	4	1	2.1	3.3	3.5
5	30/08/2009	Malaga	3	0	4	3	3.4	3.4	2.1
6	30/08/2009	Mallorca	2	0	3	3	1.91	3.4	4.33
7	30/08/2009	Osasuna	1	1	2	7	2.88	3.2	2.5
8	30/08/2009	Santander	1	4	3	6	2.15	3.3	3.4
9	30/08/2009	Valencia	2	0	6	1	2.25	3.3	3.2

Season 2018/2019 is transformed into the same structure as the Training set, and from now on, will be called the Validation set.

This is an example of such Validation set contains data from the 2017/2018 season of the Spanish LaLiga (only the first 10 rows and columns):

	date	HomeTeam	GoalsFH	GoalsAH	ShootsTargetFH	ShootsTargetAH	B365H	B365D	B365A
1	17/08/2018	Betis	0	3	8	4	1.66	4	5
2	17/08/2018	Girona	0	0	1	1	1.75	3.6	5
3	18/08/2018	Barcelona	3	0	9	0	1.11	10	21
4	18/08/2018	Celta	1	1	2	5	1.85	3.5	4.5
5	18/08/2018	Villarreal	1	2	7	4	2.04	3.4	3.8
6	19/08/2018	Eibar	1	2	6	6	1.66	3.75	5.5
7	19/08/2018	Real Madrid	2	0	3	1	1.2	7	13
8	19/08/2018	Vallecano	1	4	2	8	3.25	3.6	2.14
9	20/08/2018	Ath Bilbao	2	1	5	2	1.75	3.3	5.5

Season 2019/2020 is also transformed into the same structure as the Training set, and from now on, will be called the Test set.

This is an example of such Test set contains data from the 2019/2020 season of the Spanish LaLiga (only the first 10 rows and columns):

	date	HomeTeam	GoalsFH	GoalsAH	ShootsTargetFH	ShootsTargetAH	B365H	B365D	B365A
1	16/08/2019	Ath Bilbao	1	0	5	2	5.25	3.8	1.65
2	17/08/2019	Celta	1	3	4	11	4.75	4.2	1.65
3	17/08/2019	Valencia	1	1	6	3	1.66	3.75	5.5
4	17/08/2019	Mallorca	2	1	4	5	2.8	3.2	2.6
5	17/08/2019	Leganes	0	1	2	2	2	3.2	4.2
6	17/08/2019	Villarreal	4	4	7	7	1.6	3.8	6.5
7	18/08/2019	Alaves	1	0	2	4	2.15	3.2	3.6
8	18/08/2019	Espanol	0	2	2	4	3.2	3.3	2.3
9	18/08/2019	Betis	1	2	4	3	1.66	3.75	5.5

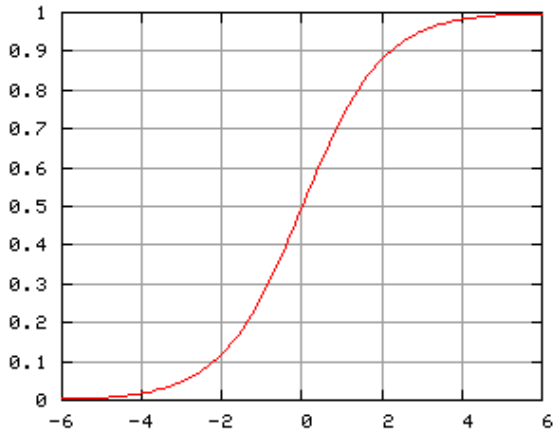
Actually, as you'll read in the algorithm part, the only features we will eventually use are the B365H, B365D, B365A, cumG and FTR:

- B365H – stands for the final ratio given to the winning of the Home team in Bet365 betting website, as explained in the bet line part.
- B365D – stands for the final ratio given to the Draw in Bet365 betting website, as explained in the bet line part.
- B365A – stands for the final ratio given to the winning of the Away team in Bet365 betting website, as explained in the bet line part.
- cumG – stands for the number of games the Home Team (Away team) has played on the current season, previously to this game.
- FTR – stands for the final result of the match – either H (Home win), D (Draw) or A (Away win)

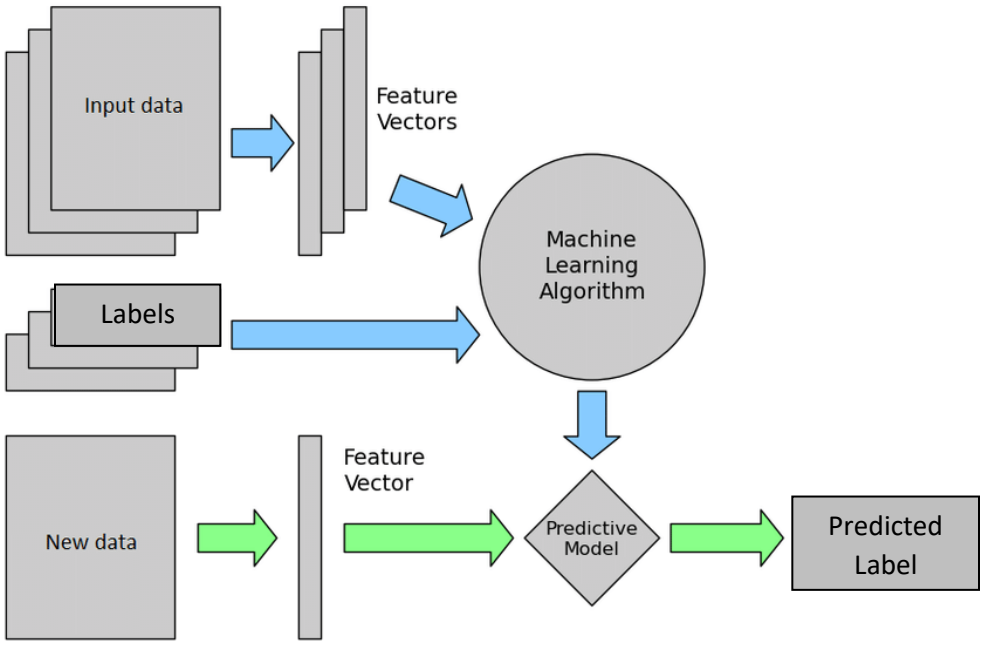
Part 4: The Algorithm

Although lots of data is used by the algorithm and that it uses a machine learning tools, the actual processing is not that complicated, and there is no need to be a machine learning expert to understand it. The algorithm is based on two things. First, finding matches with higher predicted probabilities than implied probabilities. Second, the machine learning algorithm that return the predicted probabilities, which is based on an algorithm called Logistic Regression.

Logistic Regression, also known as Logit Regression or Logit Model, is a mathematical model used in statistics and machine learning to estimate (guess) the probability of an event occurring, given some previous data. The ordinary Logistic Regression works with binary data, where either the event happens (1) or the event does not happen (0) – the most recent and actual example for this are the COVID-19 tests, that based on medical information and previous data have to estimate the probability of the patient to be positive. Logistic Regression uses the logistic function to find a model that fits with the data points. The function gives an 'S' shaped curve to model the data. The curve is restricted between 0 and 1, so it is easy to apply when the missing label is binary. In our case, the missing label in not binary (Home, Draw, Away) but we can use a similar, more general, algorithm that predicts multi-labeled data. You can read more about Logistic Regression here - [Logistic Regression].



Generally, a machine learning algorithm is developed by learning a big amount of former data, trying to reach some conclusions about future data. The input data is transformed into feature vectors whom the classic machine learning algorithms can learn together with their known labels. The machine learning algorithms develop a predictive model. And finally, after the predictive model is created, given new data, it can easily transform into feature vector that the predictive model gets as an input, and outputs the predicted label of the new data.



The transformation of the input data into feature vectors have been widely discussed on the Data Processing part. Notice that the labels are the different full-time results: Home win, Draw and Away win.

As I said, the machine learning algorithm I chose to use is Logistic Regression. In python, there is an amazing library called scikit-learn that helps you perform the machine learning process on your input data very easily. You can read more about the Logistic Regression implantation here - [scikit-learn Logistic Regression].

At the beginning, I tried using different features like shoots, shoots on target, fouls, corners, attacking and defensive power and many more, but the result was always the same… No matter what features I used, it was never better than simply use the 3 simple features: B365H, B365D, B365A. My explanation for this, is that the betting sites like Bet365 have such a strong algorithm that set the 3-way line, which probably include all of the features I tried and many more, that the B365H, B365D, B365A are the best combination I could find.

It reminds of this joke:

An engineer and a mathematician were shown into a kitchen, given an empty pan, and told to boil a pint of water. They both filled the pan with water, put it on the stove, and boiled it.

The next day they were shown into the kitchen again, given a pan full of water, and told to boil a pint of water.

- *The engineer took the pan, put it on the stove, and boiled it.*
- *The mathematician took the pan and emptied it, thereby reducing it to a previously solved problem*

So, how have I predicted the probabilities? I used the probabilities that another algorithm has already predicted :)

I used the predicted probabilities for finding those positive distances we mentioned. Finally, by narrowing down the number of matches the algorithm bets on, based on those distances and original winning ratios from the 3-way bet line, I finally found positive expected utility algorithms.

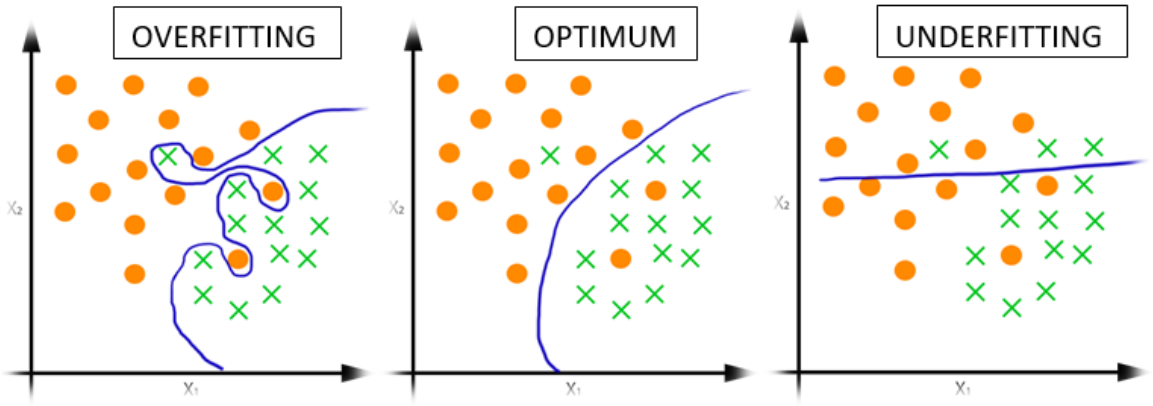
In most of the cases, the elimination of games the algorithm bets on, converged to matches with very high winning ratios, aka surprises, basically because each win on these kinds of bets is massive compared to the losses.

Part 5: Results Evaluation

As mentioned before, the objective function I tried to maximize is the total profit the algorithm makes during a football season, by betting on different matches.

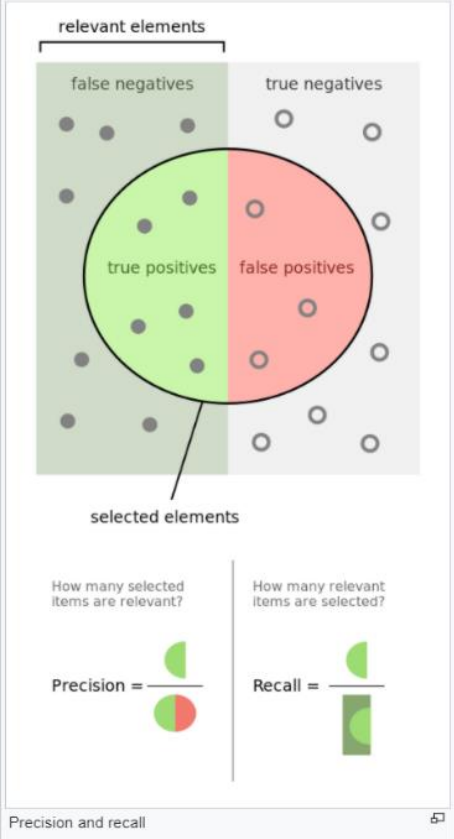
In the Algorithm part I explained how I developed different algorithms that find 'surprises' either for Home win, Draw or Away win in a certain league. Each of these algorithms have been first evaluated on the Validation set, improved over time, and then evaluated on the Test set. If after achieving good results on the Validation set, an algorithm also achieved good, and similar results on the Test set, the algorithm is ready.

As explained in the algorithm part, there are a lot of variables I changed until I reached some satisfying results. Because the results are tested over one season and each algorithm bets on a very small number of games (surprises), we are risking on overfitting – an algorithm which based on very specific, not general, data in the Test set, and after some time you can see bad results on other Test sets. On the other way, it is obviously needed that the algorithms perform well on the initial Test set, because by trying to generalize the algorithm too much, we wouldn't make smart algorithms at all. Therefore, we are looking for some optimal solutions that will work well over time, on different test sets (football seasons). The most basic way of achieving that is to use two different test sets while testing the algorithm. One will be called the Validation set and the other one will be called the Test set. The initial testing is done on the Validation set, but the final approve of the algorithm strength is depends on its performance on the Test set.



Checking the power of these algorithms will be done by reporting different measures, both on the Validation Set and the Test Set:

- Precision – we would like to know how good the algorithm did on the bets he places. The best measure for this is precision, which measures the percentage of the bets the algorithm sent that were actually winning. Another possible measure is Recall which measures how many of the Home wins / Draws / Away wins have been found by the relevant algorithm, though it is not that interesting on our case since only a very small number of bets are sent by each of our betters (only surprises).



Precision and recall

- Num bets – The number of bets the algorithm placed during the football season.

- Profit – this is the most obvious measure. The profit of an algorithm is simply the final utility after placing all the bets in the football season.

- Return of Investment - Return of Investment (ROI) is a performance measure used to evaluate the efficiency of an investment or compare the efficiency of a number of different investments. ROI tries to directly measure the amount of return on a particular investment, relative to the investment's cost. To calculate ROI, the benefit (or return) of an investment is divided by the cost of the investment. The result is expressed as a percentage or a ratio. We are interested on both the average ROI which calculated from the ROIs after every bet and the final ROI after the entire football season is ended.

I'll give you as example those measures regarding the three algorithms, I developed for the French football league (Ligue 1), given that any bet was 1$:

	Validation Set (2018/2019)					Test Set (2019/2020)				
	Num bets	Precision	Profit	Avg ROI	Final ROI	Num bets	Precision	Profit	Avg ROI	Final ROI
Home Wins Predictive Model	5	100%	17.08$	3.545	3.41	2	50%	2.5$	0.125	1.25
Draws Predictive Model	7	85.7%	15.5$	2.477	2.214	4	75%	7.25$	2.203	1.812
Away Wins Predictive Model	8	37.5%	8.25$	1.15	1.031	9	44.4%	11.95$	1.443	1.327

As expected, the algorithms suffer from a small problem of overfitting, and still in terms of economics and investments, those numbers are outstanding! The algorithms have proved their power and consistency during the season.

Looks like I have a goose that lay golden eggs in my hands.

Part 6: Warning

I have responsibility to warn you that start gambling might be harmful to your life, and there are some matters you need to consider before you start placing bets.

It's important to know that gambling problems can happen to anyone, from any walk of life. Your gambling goes from a fun, harmless activity to an unhealthy obsession with serious consequences. Whether you bet on sports, scratch cards, roulette, poker, or slots - in a casino, at the track, or online - a gambling problem can strain your relationships, interfere with work, and lead to a financial disaster. You may even do things you never thought you would, like running up huge debts or even stealing money to gamble.

Pathological gambling, aka Gambling addiction, is an impulse-control disorder. A compulsive gambler can't control the impulse to gamble, even when it has negative consequences for him or his loved ones. He will keep gambling whether he is up or down, broke or flush, and he will keep gambling regardless of the consequences - even when he is aware to the fact that the odds are against him and he can't afford another lost.

Of course, you can also have a gambling problem without being totally out of control. A gambling problem is any gambling behavior that disrupts your life. If you're preoccupied with gambling, spending more and more time and money on it, chasing your losses, or keep gambling despite serious consequences in your life, unfortunately, you are suffering from a gambling problem and you must take care of that.

A gambling addiction or problem is also associated with other behavior or mood disorders. Many problem gamblers also suffer with substance abuse issues, unmanaged ADHD, stress, depression, anxiety, or bipolar disorder.

If you feel that you already felt into that dark hole, it's important that you know that although it may feel like you're powerless to stop gambling, there are plenty of things you can do to overcome the problem, repair your relationships and finances, and finally regain control of your life – I recommend you reading more about that here: Gambling Addiction

Part 7: Summary

To Sum up, I am glad I had the opportunity to combine 2 of my biggest loves, and work on a project that actually interest me a lot. I have big confidence in the power of the algorithms I developed, and looking forward to check them in the upcoming European seasons.

I remind you again the dark sides of gambling, and ask you the following questions:

- Do you have spare money to spend on the bets?
- Have you found a real, strong, proved, strategy to bet on games, and not just follow your gut-feeling?
- Do you have spare time to spend on betting over time?
- Do you feel like you have a strong personality that will let you quit if you see you are losing money over time?

If the answer to one or more of these questions is no, **you should not, and I repeat, you should not, start gambling**.

Back to me. I have no doubt that this project taught me a lot about Data Science and the statistics behind this huge sport called football.

After finishing this project, I came up with even more ideas of improving the algorithm I developed, finding new algorithms, use even more data and machine learning tools, and finally of course, implement a similar process on even more leagues and sports, like basketball in the NBA league.

If you have any new ideas to improve the algorithms, or want a further explanation on the project or the code behind it, you are more than welcome to contact me at: weissman.gil@gmail.com

Part 8: Bibliography

1. Google Predict football matches
2. bet365
3. Winner
4. zero-sum game
5. Zero-sum game solver
6. Football Datasets Website
7. Logistic Regression
8. scikit-learn Logistic Regression
9. Gambling Addiction