香港中文大學

計算機科學及工程學系

Department of Computer Science and Engineering,

The Chinese University of Hong Kong

FYP Report (First Term)

Exploiting Betting Odds using Machine Learning

Written By

WONG Wing Keung (1155093416)

WONG Ching Yeung Wallace (1155093534)

Supervised By

Prof. Michael R. Lyu

Contents

Tables

Abstract

In this project, we would apply machine learning to forecast sport events and evaluate the performances by simulate betting against bookmakers. Unlike most research projects that use performance metrics of players, teams etc. for prediction, we make use of the betting odds to forecast sport events.

In this term, we developed a method that predicts probabilities of sport events in sport betting markets. We tested the method on the horse racing in Hong Kong and the method was shown to be profitable when betting against bookmakers.

Acknowledgements

We would like to thank our supervisor Prof. Michael R. Lyu and advisor
Mr. Edward Yau for their guidance and feedback.

In addition, we appreciate the Department of Computer Science and Engineering,
The Chinese University of Hong Kong for offering the required computing resources.

Disclaimer

According to the Gambling Ordinance, Chapter 148, Laws of Hong Kong, all gambling activities are illegal except those authorized by the Government.

Results included in this report are simulations. No participation in any forms of gambling is involved. We have no intention of promoting or facilitating illegal betting or bookmaking.

1. Introduction

1.1 Motivation

Sport betting markets are getting more popular nowadays. There is an increasing number of online bookmakers offering betting markets for uncertain events that occur in sports games. From 2009 to 2016, the market size of the global online gambling market doubled gradually from 20 billion USD to 40 billion USD [1]. At the same time, machine learning has been shown to be successful in applying to multiple fields and industries in recent years. We want to explore if machine learning can beat the bookmakers in sport betting.

1.2 Background

1.2.1. Types of Betting

There are 2 types of betting system in general – "pari-mutuel betting" and "fixed-odds betting".

In "pari-mutuel betting", bets are placed into a "pool", which is operated by a bookmaker. The bookmaker will deduct a portion of bets from the pool as commission fees. After that, winners will share the remaining amount of money in the pool in proportion to their winning stakes.

In "fixed-odds betting", bettors will bet for the odds which are offered by bookmakers. Although odds may be adjusted from time to time until closing, the payout is based on the odds at the time that the bet is accepted. Odd changes may due to the bettors' betting activities. Pinnacle, an online bookmaker which offers almost the highest average odds among all major bookmakers [2], claimed that they will make use of the betting activities of their "sharp" bettors to correct their odds [3].

The Hong Kong Jockey Club, the only legal bookmaker in Hong Kong, accepts bets for local horse racing and soccer matches. "pari-mutuel betting" and "fixed-odds betting" systems are used for horse racing and soccer matches respectively.

1.2.2. Limitation on "pari-mutuel betting"

Due to the nature of "pari-mutuel betting", bettors are unable to know their payouts exactly until the pool is closed. In horse racing, although the Hong Kong Jockey Club provide "odds" while accepting bets, the "displaying odds" are calculated based on the pool at that moment. It is subject to change when others' bets are going into the pool afterwards.

In order to approximate the final payouts before placing bets, William Benter, a well-known bettor in Hong Kong horse racing market, suggested placing bets as late as possible [4]. The idea is that the sooner you place your bets, the "displaying odds" at that time will be closer to the final one.

However, from our observation, the last "displaying odds" that bettors can see just before the pool is closed, are still very different from the final one. Therefore, in this project, we will only focus on "fixed-odds betting", which allows bettors to know their payouts before placing the bets.

1.3 Objective

The objective is to develop betting-oriented methodologies that use machine learning to exploit the fixed-odds betting markets. Methodologies will be evaluated on horse racing and soccer betting markets.

The project is expected to complete in 2 terms. The following are the objectives for the first term:

- Develop a profitable method that use machine learning to forecast probabilities of sport events.
- Test the proposed method in Hong Kong horse racing.

2. Related Work

A regular approach to sport events forecasting was to build predictive models using statistical modelling. As machine learning becomes more popular in recent years, people started to adopt it for the problem. There are mainly 3 types of approaches.

2.1 Performance-based Forecast

Many studies adapted this approach to sport events forecasting by modelling event outcomes from historical performance metrics.

For soccer, Capobianco et al. used possession statistics, the number of recovered, the average speed of a team etc. to build classifiers to predict the winner and total goal in a game. Their best classifiers achieved precisions of at least 0.8 in all predictions. [5]. Nazim et al. used other metrics – number of shots, number of corners, number of fouls, number of cards etc. to train a Bayesian classifier to predict the match winner in English Premier League. The classifier was claimed to be achieved 75.09% of accuracy [6]. Besides from measurable metrics, Baboota et al. included team ratings from FIFA, a well-known soccer video game, as features in their probabilistic machine learning models. Although their best model achieved promising Ranked Probability Scores of 0.2156, when coming to predict match winners in English Premier League, their best model was not able to outperform bookmakers' predictions [7].

For horse racing, Khanteymoori et al. used horse weight, race distance, track condition, weather etc. as features of their neuron network model to predict the final positions of horses in Aqueduct Racing, a horse racing event in New York. The average accuracy on the training set was claimed to be 77%. However, the paper did not include performance evaluation on unseen dataset [8]. Previous FYP students Liu, Cheung and Lau used similar features to build predictive classifiers for horse racing in Hong Kong. All of them failed to generate profits in general but selectively under some conditions when placing bets on Hong Kong Jockey Club [9] [10].

2.2 Odds-based Forecast

Odds-based Forecast makes use of betting odds in the sport betting market for prediction. Nowadays, many bookmakers accept bets online. Betting odds are available online and can be accessed easily. Some researchers started to look at the importance of the betting odds to this problem.

2.2.1. Efficiency of Fixed-Odds Betting Market

Market Efficiency is a concept in financial economics that measures the correlation of prices and information in a market [11]. A betting market is described as efficient if the betting odds closely reflect the true probabilities. In other words, no profitable strategy exists if the market is efficient enough.

Efficiency of different sport betting markets has been evaluated by quite a lot of studies. It is still controversial to conclude the betting markets are fully efficient or inefficient. Multiple bookmakers may offer the same market for a single sport event, for example, match winner in a soccer game. Recent studies suggest the betting markets are not inefficient in general [12] [13] [14] [15], but inefficiency is observed [12] [15].

2.2.2 Opening Odds and Closing Odds

In 1998, Gandar et al. studied the betting odds in NBA betting market and observed the same phenomenon that the closing odds were more accurate than the opening odds. They suggested odds changes in the market were not simply noise but may due to betting activities from informed bettors, who got more information about the matches than the general public and the bookmakers [16]. A recent study in 2016 investigated betting odds in college basketball games and further suggested that odds movements were made to prevent informed bettors taking too many advantages from the original odds [17].

2.2.3 Using Betting Odds to Forecast Sport Events

Although betting odds (especially the closing odds) are shown to be accurate by various studies, there are limited studies regarding turning the betting odds into predictors. There are mainly 2 ways to apply the betting odds to the problem.

2.2.3.1 Converting Odds to Probabilities

For an outcome, lower odds often imply a higher probability for it to occur. Therefore, the probability is inversely proportional to the betting odds. If the sum of reciprocals of all odds is equal to 1, then the odds-implied probability of an outcome will be given by:

$$\text{implied probability of an outcome} = 1 / (\text{odds of that outcome})$$

However, bookmakers set “margins” in their odds for making profits. The sum of reciprocals of odds is larger than 1. In this case, the above relation will not hold as the sum of implied probabilities for all outcomes will be larger than 1, which does not make sense. Therefore, adjustments have to be made to the implied probabilities in order to force the probabilities sum to 1.

One simple way, known as the Normalization Method, is to normalize each implied probability by the sum of reciprocals of odds.

$$\text{adjusted implied probability of an outcome} = \frac{\text{implied probability}}{\text{sum of reciprocals of odds}}$$

Bookmakers often set higher margins for high odds in order to reduce their risks. Thus, the Normalization Method may overestimate the probabilities of the underdogs. Clarke suggested another way, known as “The Power Method”, for adjustment. The implied probabilities are raised to a constant power, such that they sum to 1 [18]:

$$\sum_i^n \text{ (implied probability of outcome } i)^c = 1 \text{ , where c is a constant}$$

$$\text{adjusted implied probability of an outcome} = (\text{implied probability})^c$$

The Power Method reduces lower implied probabilities more. In order words, higher odds will be penalized more. It has been shown to be more accurate than the Normalization Method by Clarke in a later study [19]. However, there is no direct solution exists for the constant *c*. Computing *c* requires iterative methods or optimizers in general.

Although there are other conversion methods - “The Shin Method” [20] [21] and “The Additive Method” [22], they may not able to produce implied probabilities that sum to 1 exactly.

Finally, to output forecasts, simply convert the odds into probabilities using a method described above.

2.2.3.2 Odds-Based Classifier

At the moment, there are only a few studies attempted to use machine learning approaches to deal with the betting odds. Grekow et al. used the odds change data from a single bookmaker as features to build classifiers. Their best classifier achieved 70.30% of accuracy when predicting match winners (ignoring draw) in soccer games [23].

2.3 Odds-and-Performance-based Forecast

This type of approach uses both betting odds and performance metrics to produce forecasts. It is a common approach to the problem. Engin et al. used performance metrics of teams from the last 6 games and betting odds to build KNN classifiers to predict match winners in English Premier League. The best accuracy achieved was 57.52% [24].

There are lots of other attempts using this approach. In general, numerous of them are the mix of the details in 2.1 Performance-based Forecast and 2.2 Odds-based Forecast. And thus, their details will not be covered here.

3. Methodology

3.1. Overview

After a careful study, we decided to develop a method that produces Odds-Based Forecast. This is because performance metrics may not be available in every sport. For example, for a soccer game, there is no way to determine the number of attacks, the number of defenses etc. unless the game is ended. Some previous studies would use performance metrics from the past few games for prediction. This may result in inaccuracy, as past performances depend on performances of the opponents and are no guarantee of future results. In contrasts, betting odds are accessible for every match before the kick-off time.

The goal of this project is to develop a betting-oriented method. From authors' perspectives, knowing the chances and advantages is the key to success. Instead of classifiers, probabilistic models will be more suitable for this problem.

3.2 Betting Strategy

After having the predicted probabilities from our models, we can easily compute the expectations of each outcome in a market. In probability theory, betting for outcomes with negative expectations will result in bankruptcy in the long run. Therefore, our strategy should only bet for those with positive expectations. Besides, a wagering strategy that can produce the maximum return is needed. In gambling theory, there is a well-known formula that relates betting odds and probabilities – Kelly Formula [25].

3.2.1 Kelly Formula

Kelly Formula is used to calculate the optimal fraction of current capital that should be placed, such that the expected geometric growth rate can be maximized, given the odds and probability of winning are known in a game.

The most common version of Kelly Formula K is as follows:

$$K(\sigma,\ p) = \frac{p\sigma - 1}{\sigma - 1}$$

, where p is the probability of winning and σ is the odds offered.

Here is the deviation:

Suppose p is the probability of winning, σ is the odds offered, k is the ratio of the capital to the bet size, the overall rate of return (E) after n (large enough) repeated betting will be:

$$E = (1 + k(\sigma - 1))^{np}(1 - k)^{n(1-p)}$$

$$\log E = np \log(1 + k(\sigma - 1)) + (n - np) \log(1 - k)$$

The k that maximizes $\log E$ can be found by solving:

$$\frac{d \log E}{dk} = \frac{np\sigma - n + nk - nk\sigma}{(1 - k)(k(\sigma - 1) + 1)} = 0$$

$$k = \frac{p\sigma - 1}{\sigma - 1}$$

3.2.2 Kelly Betting

The betting strategy that utilizes Kelly Formula is known as Kelly Betting. Kelly Betting requires an initial capital to start. Whenever we bet, we use Kelly Formula to compute the optimal wager:

$$\text{optimal wager} = \text{current capital} \times K(\sigma,\ p)$$

If the Kelly Formula gives a negative result, it means the expectation is negative and we should avoid placing bets on that outcome. It can be easily shown:

When the expectation is negative, the "fair odds" is larger than the one offered by the bookmaker. And thus, $\frac{1}{p} > \sigma \quad \Rightarrow p\sigma < 1 \quad \Rightarrow p\sigma - 1 < 0 \quad \Rightarrow K(\sigma,\ p) < 0$

3.3 Model Construction

3.3.1 Data Preparation

Here we will introduce the general procedure of forming the dataset.

Let us consider a betting market of a match or a sport game.

We convert the odds into odds-implied probabilities by a predefined method described in 2.2.3.1 Converting Odds to Probabilities.

After that, for every possible outcome in the market, we create a record of the dataset which contains the features and the label. The label is a Boolean value indicates whether it is the final outcome of the game. And the features are created by considering a period of average odds-implied probabilities.

Details are as follows:

Suppose there are n possible outcomes, k bookmakers (b) offering odds for those n outcomes, and odds in a period t will be considered.

Let the odds offered by bookmaker b_g at time t_h for the ith outcome be $o(b_g, i, t_h)$.

Let the odds-implied probabilities from bookmaker g of ith outcome at time t_h be $P(b_g, i, t_h)$.

Let $P_{\text{avg}}(i, t_h) = \frac{1}{k} \sum_j^k P(b_j, i, t_h)$, $C(i, t_h) = \frac{P_{\text{avg}}(i, t_1) - P_{\text{avg}}(i, t_h)}{P_{\text{avg}}(i, t_h)}$, $S_i = \{ C(i, t_2), C(i, t_3), C(i, t_4), \ldots C(i, t_m) \}$

For every outcome i, we first compute the followings as features to capture the general changes in odds:

�������

����

}

Next, d degree ���� The coefficients $\{ c_1, c_2, \ldots c_{d+1} \}$ in the polynomial regression fit will be included in the features.

Therefore, there will be 10 + d features for outcome i .

3.3.2 Model Structure

The model structure is expected to be a sequential neutron network model. The only requirement is that the activation function in the last layer must be a Sigmoid function or any other functions that produce values ranging from 0 – 1, as we are going to construct a probabilistic model.

3.3.3 Training

3.3.3.1 Loss Function

A common loss function - Binary Cross Entropy will be chosen to be the loss function. Here we will show that Binary Cross Entropy is the optimal loss function to use with Kelly Betting.

Recall that, the Kelly Formula is given by

$$K(\sigma_i,\ p_i) = \frac{p_i\sigma_i - 1}{\sigma_i - 1}$$

, where p_i is the probability of winning (predicted) and σ_i is the odds offered in an outcome *i* in an event.

Suppose we have *n* outcomes in total, and y_i is the label of the outcome i (whether it is the final outcome or not), σ_i is chosen to be some closing odds available in the market.

The rate of return (V') after performing Kelly Betting in these *n* outcomes:

$$V' = \prod_i^n \ (1 + \max(0, K(\sigma_i,\ p_i)) \times (\sigma_i\ - 1))^{y_i} (\ 1\ - \max(0, K(\sigma_i,\ p_i)))^{1-y_i}$$

Maximizing V' is no different from maximizing V and $\log(V)$:

$$V = \prod_i^n \ (1 + K(\sigma_i,\ p_i) \times (\sigma_i\ - 1))^{y_i} (\ 1\ - K(\sigma_i,\ p_i))^{1-y_i}$$

$$V = \prod_i^n \ (p_i\sigma_i)^{y_i} (\frac{\sigma_i - p_i\sigma_i}{\sigma_i - 1})^{1-y_i}$$

$$\log(V) = \sum_i^n \ y_i\log(p_i\sigma_i) + (1 - y_i)\log(\frac{\sigma_i - p_i\sigma_i}{\sigma_i - 1})$$

The partial derivative of $\log(V)$ with respect to p_a where $i \leq a \leq n$ is given by:

$$\frac{\partial \log(V)}{\partial p_a} = \frac{y_a - p_a}{p_a(1 - p_a)}$$

On the other hand, the Binary Cross Entropy (BCE) is given by

$$\text{BCE} = \frac{1}{n}\sum_{i}^{n} \; y_i \log{(p_i)} + (1 - y_i)\log{(1 - p_i)}$$

The partial derivative of BCE with respect to p_a where $i \leq a \leq n$ is given by:

$$\frac{\partial \text{ BCE}}{\partial p_a} = \frac{y_a - p_a}{p_a(1 - p_a)} = \frac{\partial \log(V)}{p_a}$$

This implies that optimizing the Binary Cross Entropy is no different from optimizing $\log(V)$, which is a measure of the rate of return in Kelly Betting. Therefore, Binary Cross Entropy is the optimal loss function to use when Kelly Betting is chosen to be the betting strategy.

3.3.3.2 Early Stopping

Overfitting will very likely to cause bankrupt in Kelly Betting, as the bet size is highly related to the predicted probability. Overestimation should be avoided as possible. Therefore, Early Stopping will be used during training to reduce over-fitting. To train a model, we first shuffle the whole training set. The first half of data will be used in training and the second half will be used to monitor the loss continuously. Training will be stopped if the monitored loss shows no improvement in the last 500 epochs.

3.4 Ensemble Forecast

Each of the trained models carries its own hypothesis. A different model will be produced when we run the training again, especially under Early Stopping (mentioned in *3.3.3.2 Early Stopping*), where a different subset of the training set will be used for training every time. As a result, the performances of trained models can be different.

Instead of training 1 model, multiple models are trained and grouped to form an ensemble model. The output of the ensemble model will be the average of outputs from its ensemble members.

4. Application in Hong Kong Horse Racing

In Hong Kong, there are nearly 700 horse races conducted at Sha Tin Racecourse and Happy Valley Racecourse per year. We applied the proposed method to predict the race winner in Hong Kong horse racing.

4.1 Odds Data

Offshore bookmakers have offered fixed-odds markets for horse racing in Hong Kong. We collected the odds data of races in 2017/01/01 – 2019/10/01 from 15 distinct bookmakers. **Table 1** displayed a list of targeted bookmakers which offer markets for Hong Kong horse racing by year.

Year	Count	Bookmakers
2019	12	Bet365, Bet Easy, Betstar, Bluebet, Bookmaker, Ladbrokes, Neds, Pointsbet, Sportsbet, Sportsbetting, Topbetta, Unibet
2018	13	Bet365, Bet Easy, Betstar, Bluebet, Bookmaker, Ladbrokes, Neds, Pointsbet, Sportsbet, Sportsbetting, Topbetta, Ubet, Unibet
2017	9	Bet365, Betstar, Bookmaker, Ladbrokes, Neds, Pointsbet, Sportsbet, Topbetta, Unibet

Table 1 List of Bookmakers offer markets for Hong Kong horse racing

4.2 Data Preparation

Odds in 2017 – 2018 were used to form the training set, and the data in 2019/01/01 – 2019/10/01 were used for performance evaluations. We followed the procedures described in *3.3.1 Data Preparation* to prepare the data.

4.2.1 Experiments

For all sets of models, odds-to-probabilities conversion method was chosen to be the normalization method and odds sampling interval was set to be 1 minute.

The abbreviation X-Y-Zdeg is used to describe an ensemble model with odds considering period X minute before closing – Y minute before closing and Z degree of polynomial regression.

4.3 Model

A sequential model with 14 fully connected layers was used. Except for the last layer, all layers were using ReLu as activation. Sigmoid was used as activation in the last layer. Besides, each ensemble model was filled up with 1000 ensemble members.

4.4 Betting Simulation

Models would be tested using the odds data in 2019/01/01 – 2019/10/01. For every horse picked, the highest closing odds offered among different bookmakers will be chosen for betting. For all simulations, the initial capitals were set to $10000.

It is possible that multiple horses will be picked in a single race. The results are revealed simultaneously after the race is ended. The optimal bet that should be placed on a horse will also depend on that of others. Modification to the original version of Kelly Formula in *3.2.1 Kelly Formula* is needed for this case. However, for simplicity, the betting simulations below assumed that the results of every outcome are known after the bets are placed. This makes each betting becomes independent and the original Kelly Formula can be applied.

4.5 Simulation Results

Model	Return in Kelly Betting
0-9-6deg	9061
0-19-6deg	13311
0-19-10deg	15125
0-19-16deg	107
0-29-6deg	29254
0-29-10deg	32271
0-29-16deg	6302
0-49-6deg	17176
0-49-10deg	17434
0-49-16deg	16238
0-79-6deg	9529
0-79-10deg	9791
0-79-16deg	12201
0-119-6deg	5850
0-119-10deg	6602
0-119-16deg	6491

Table 2 All Simulation Results

5. Conclusion

In this report, we proposed a method that forecasts probabilities of sport events with machine learning. Our method makes use of the changes in betting odds only and does not require any performance metrics of a game. We tested the method on the horse racing in Hong Kong and the method was shown to be profitable when betting against bookmakers.

6. Reference

[1] C. Gough, "Sports Betting and Gambling Market/Industry - Statistics & Facts," Statista, 29 3 2019. [Online]. Available: https://www.statista.com/topics/1740/sports-betting/.

[2] "Odds Quality - Bookmaker Payout Ratings," OddsPortal, [Online]. Available: https://www.oddsportal.com/odds-quality/. [Accessed 29 10 2019].

[3] "Winners are Welcome - Successful players sharpen our odds," Pinnacle, [Online]. Available: https://www.pinnacle.com/en/promotions/winners-welcome.

[4] ""What Are My Odds?" - William Benter ICCM 2004," [Online]. Available: https://youtu.be/YOVrZrJ-wtc?t=2097. [Accessed 2019 10 29].

[5] Giovanni Capobianco, Umberto Di Giacomo, Francesco Mercaldo, Vittoria Nardone, Antonella Santone, "Can Machine Learning Predict Soccer Match Results?," ICCART, 2019.

[6] Nazim Razali, Aida Mustapha, Faiz Ahmad Yatim, Ruhaya Ab Aziz, "Predicting Football Matches Results using Bayesian Networks for English Premier League (EPL)," IOPscience, 2017.

[7] Rahul Baboota, Harleen Kaur, "Predictive analysis and modelling football results using machine learning approach for English Premier League," International Journal of Forecasting, 2018.

[8] Ali Reza Khanteymoori, Elnaz Davoodi, "Horse Racing Prediction Using Artificial Neural Networks," Recent Advances in Neural Networks, Fuzzy Systems & Evolutionary Computing, 2010.

[9] Y. LIU, "Predicting Horse Racing Result with Machine Learning," 2018.

[10] LAU Ming Hei, CHENG Tsz Tung, "Predicting Horse Racing Result using Tensorflow," 2016.

[11] Elroy Dimson, Massoud Mussavian, "MARKET EFFICIENCY," in *THE CURRENT STATE OF BUSINESS DISCIPLINES*, SPELLBOUND PUBLICATIONS, 2000, pp. 959 - 970.

[12] Guy Elaad, J. James Reade, Carl Singleton, "Information, Prices and Efficiency in An Online Betting Market," Applied Economics Letters, 2019.

[13] Štefan Lyócsa, Tomáš Výrost, "To bet or not to bet: a reality check for tennis betting market efficiency," Applied Economics, 2017.

[14] Luca Rebeggiani, Johannes Gross, "Chance or Ability? The Efficiency of the Football Betting Market Revisited," IASE Conference, 2018.

[15] Giovanni Angelini, Luca De Angelis, "Efficiency of online football betting markets," SSRN Electronic Journal, 2017.

[16] John M. Gandar, William H. Dare, Craig R. Brown, Richard A. Zuber, "Informed Traders and Price Variations in the Betting Market for Professional Basketball Games," THE JOURNAL OF FINANCE, 1998.

[17] Kevin Krieger, Andy Fodor, "Price movements and the prevalence of informed traders: The case of line movement in college basketball," Journal of Economics and Business, 2013.

[18] S. R. Clarke, "ADJUSTING TRUE ODDS TO ALLOW FOR VIGORISH," Proceedings of the 13th Australasian Conference on Mathematics and Computers in Sport, 2016.

[19] S. R. Clarke, "Adjusting Bookmaker's Odds to Allow for Overround," American Journal of Sports Science, 2017.

[20] H. S. Shin, "Prices Of State Contingent Claims With Insider Traders, And The Favourite-Longshot Bias," in *The Economic Journal*, 1992, pp. 426-435.

[21] H. S. Shin, "Measuring the Incidence of Insider Trading in a Market for State-Contingent Claims," in *The Economic Journal*, 1993, pp. 1141-1153.

[22] M Viney, A Bedford, E Kondo, *Incorporating over-round into in-play markov chain models in tennis,* Las Vegas, USA: 15th International Conference on Gambling & Risk Taking, 2013.

[23] Jacek Grekow, Karol Odachowski, "Using Bookmaker Odds to Predict the Final Result of Football Matches," in *Knowledge Engineering, Machine Learning and Lattice Computing with Applications*, Springer, Berlin, Heidelberg, 2012, pp. 196-205.

[24] Engin Esme, Mustafa Servet Kiran, "Prediction of Football Match Outcomes Based on Bookmaker Odds by Using k-Nearest Neighbor Algorithm," *International Journal of Machine Learning and Computing,* vol. 8, 2018.

[25] J. L. Kelly, The Bell System Technical Journal, 1956.

[26] Joseph Buchdahl, "The maths behind Pinnacle's “winners welcome” policy," 20 7 2016. [Online]. Available: https://www.pinnacle.com/en/betting-articles/educational/why-pinnacle-doesnt-close-or-limit-accounts. [Accessed 29 10 2019].

[27] W. Benter, "Computer Based Horse Race Handicapping and Wagering Systems: A Report," 1994.