CHANCER: A MOBILE APP SOLUTION TO PROBLEM GAMBLING IN GHANA

Kofi Appiah
Ashesi University
Berekuso, Accra
kofi.appiah@ashesi.edu.gh

Dr Elena Rosca
Ashesi University
Berekuso, Accra
erosca@aucampus.onmicrosoft.com

***Summary* - Four in ten Ghanaians are prone to gambling addiction, fueled mainly by the hope to alleviate poverty. In particular, sports betting has been recorded as the predominant gambling activity. Due to the dramatic increase in the availability of smartphones in the Ghanaian communities (over 50% of Ghanaians have a smartphone), we proposed the development of a mobile app, Chancer. This will be a platform that incorporates a machine learning (ML) model to increase the odds of winning (above the chance 33%) and allow gamblers to manage their bets to prevent their financial betting from reaching an untenable state. Currently, the best football prediction model is a Naïve-Bayes model with an accuracy of 51.6%. Furthermore, research has demonstrated that mobile apps can address many health behaviors. Therefore, further research must be performed to profile the Ghanian gambler to determine the most effective methods in the Ghanaian context. We hypothesize that the deployment of the Chancer mobile app will lead to a decrease in problem gambling behavior if used by Ghanaians suffering from problem gambling and gambling addiction.**

***Keywords* – Machine Learning, Problem gambling.**

I - INTRODUCTION

Gambling has become extremely prevalent in Ghana. As mobile technology proliferation rises in Ghana there are more avenues for gamblers to gamble increasing the frequency of gambling. As of 2020, three in ten Ghanaian females that inhabit rural areas are prone to gambling addiction one less than their male counterparts. It is proven that these adolescents suffer greatly from external factors that push them into the gambling scene [1]. Finding a solution to these negative external factors such as poverty can be considered as a more permanent and longer-lasting solution to gambling addictions, but the nature of the solution to poverty requires a great deal of effort and due to the wide scope can only be solved with a great deal of government intervention. Hence, a focus on solving the issue of gambling should be focused elsewhere.

Despite gambling being viewed with negative connotations and can ruin one's life, it is against Ghana's ethical framework to deny one access to gambling games because it tramples on the right to freedom of expression. However, it is known that the harm induced by sports betting is a public health issue as it leads to individuals and potentially their relatives and dependents losing their financial standing and suffering greatly [2]. Further research evidence shows that as gambling frequency increases the attraction and addiction towards gambling rises [3]. Gambling addiction causes problem gambling - a situation when gambling gives rise to harm to the individual player, and/or to his or her family, and may extend into the community [4]. Hence, even though we need to avoid the ethical conflict, attention needs to be dedicated to aiding these vulnerable members of the community because the potential harm endured by gamblers and their acquaintances as gambling addiction progresses in an individual is significant.

Figure 1: Gambling market share in Ghana

There exist many ways that cell phones can be used to change health behavior including mobile applications (apps). Apps have the potential in addressing many health behaviors and have demonstrated considerable popularity [5]. Therefore, the solution to gambling disorder we propose is Chancer mobile app, a standalone self-help approach to dealing with gambling addiction. Chancer mobile app takes an alternative approach to solving gambling by incorporating behavorial change techniques common with addiction prevention applications with a machine learning model that predicts the outcome of football matches. The purpose of the machine learning model is to help the gamblers improve their gambling success rate and mitigate losses because research by Markham et al proves that with more losses in gambling come more self-harm [6]. Furthermore, the focus of Chancer is on football as it is the most popular sport and sports betting is the most common form of gambling in the country as shown in Figure 1 [7]. Additionally, these gamblers harm themselves by spending beyond their means, hence, by tracking the amount they spend on betting sites relative to their income they can be informed with striking psychologically informed imagery when their finances are gradually worsening to restrain themselves and prevent the situation from getting out of hand.

II – RELATED WORKS AND LITERATURE REVIEW

The first stage of the solution was to design a machine learning model that predicts the outcome of games at a good enough rate to ensure the gamblers do not make losses. Research into machine learning models for football prediction to gain an understanding of the procedures involved in building such a model and to decide the right goals the model must achieve. From the research, the most in-depth machine learning model for predicting football games was a Bayesian network model named 'pi-football' [8]. Pi-Football used four generic factors for both the home and away team: the team strength, the form, the psychological relationship between the teams and the game, and the physical state and fitness. The pi-football managed to achieve an accuracy of 37.2%. A figure of less than 85% is normally considered inadequate for a machine learning model however, considering that the model is classifying three possibilities; a home win, an away win, or a draw and these events have a random chance of 33%, a figure of 37.2% is an acceptable figure. Furthermore, the pi-football study despite the "low percentage" it was proved that gamblers still achieved a profit, albeit small, by using predictions from the pi-football model. The Bayesian network model managed to make accurately predict events that had a lower probability of occurring and went against the odds provided by the gambling firms due to the exhaustive nature of the data used in the training of the model. The ability to be relatively accurate on higher odds events was the main reason pi-football served as the framework to be used in data collection for this study.

The only issue with the Pi-football study was that it was not particularly recent. The research was conducted in 2012 before the advent of advanced analytics in football. Around 2014, the statistical models that football clubs used to track their performance beyond the normal statistics such as goals scored were more available for public use. The reveals were driven by analysts within the football clubs setting up their private analysis firms. These statistical models can be described as advanced analytics because the information derived is given a more advanced view of the matches beyond the details on the surface. Hence, using these statistics along with the factors from the pi-football model, a more accurate, more advanced model than pi-football can be developed.

The next stage of the solution was to investigate the best ways to warn sports bettors that they are in a precarious financial position, therefore should quit in that instance before the situation spirals out of control. Research findings from Block et al show that recall of antidrug advertising was associated with a lower probability of marijuana and cocaine use. However, recall of such advertising was not associated with the decision of how of the drug to use [8]. These findings substantiate that anti-drug advertisements are good for addicts and therefore, it would be useful to incorporate them in the application to help the gamblers.

III – SOFTWARE ARCHITECTURE

A. Front-end design and user functions.

The Chancer app functions in a single mode, the user mode. A user, which can be represented, for instance, by a problem gambler can perform several activities when logged into the application. (i) The user can access match predictions for games by clicking on the leagues they want to see. (ii) The user can access the financial information where they can input the bets – (the amount they bet and the amount they win from the bets) they have made over a period to practice self-monitoring and set limits on their spending. If the spending limits are exceeded then an image suggesting putting an end to gambling activities pops up.

B. Back-end design

The back-end of the Chancer app is composed of several software components. (i) The Client is the part managed by the app users. It uploads all the user's data to the database to keep track of it, and it responds to any user input such as entering the bets for the week. (ii) Data manager, which collects the football data in a real-time database, as well as stores the previous results and upcoming fixtures in the football calendar. (iii) Processing engine is at the server-side and holds the logic interface of the application, performing the machine learning model sending the predictions back to the client.

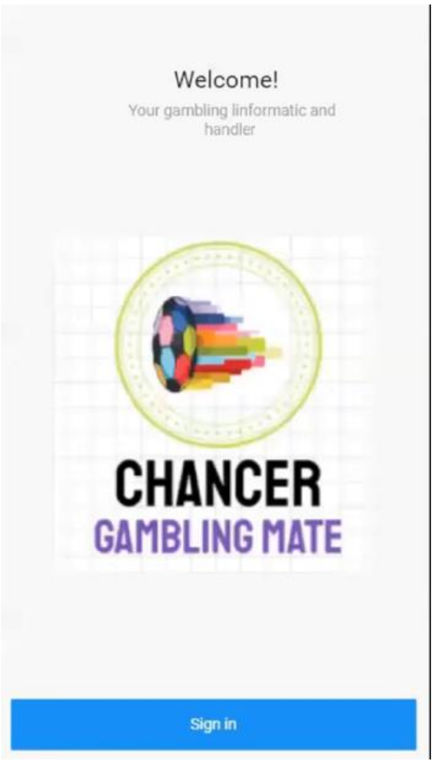

Figure 2: Front-end design of the Chancer mobile app.

IV – MACHINE LEARNING MODEL

A. Problem Description

The focus of this section is on the development of the machine learning model that is to be integrated into the Chancer app. The prediction of the result of a football game is a classification problem, as there are three discrete outcomes possible in a football match; a home win, an away win, or a tie.

B. Data collection

The data collection for input into the machine learning model was easy to obtain as the focus was on the five most-watched football leagues in the world. Due to their popularity, data on matches that take place in these leagues are easy to download from credible sources on the internet. The data points for the model are football matches and more than 7,000 games were analyzed for this study.

The statistical information used for the model are described below:

- Goals – The aim of a football match is to score a goal; the act of putting the ball in the back of the net.
- Results – The outcome of the football game is determined by which team has more goals scored in the match. There are three possibilities; home win, draw and away win.
- Expected goals – The number of goals a team is expected to score based on the location of their shots. (Collected from understat.com)
- Non-shot expected goals – the number of goals a team is expected to score based on the location, type, and proximity of their non-shooting actions. (collected from fivethirtyeight.com)
- Importance of the game – how much pressure there is on a team to win a game based on the potential consequences of the result. (collected from fivethrityeight.com)
- Sporting power index – the strength of a team calculated by the FiveThirtyEight supercomputer using Elo based system. (collected from fivethirtyeight.com)
- Distance traveled for a game – Distance covered by an away team for a game indicates the home advantage and fatigue of the away team. (collected data on location manually and had to calculate using distance information)
- Days rest between matches – The amount of rest the teams get between matches (calculated by dates between games)

C. Approach

The data collected is just on a game-by-game basis. However, predicting the result of a football match that has not occurred yet is dependent on data from previous games, hence, the data had to be processed in the Python programming language to produce data features that are segregated on a team basis and for the two teams involved in the match to be predicted. In Python programming language, the data features and their labels, the result of the football match, were tested on various machine learning algorithms and evaluated to find the most appropriate model.

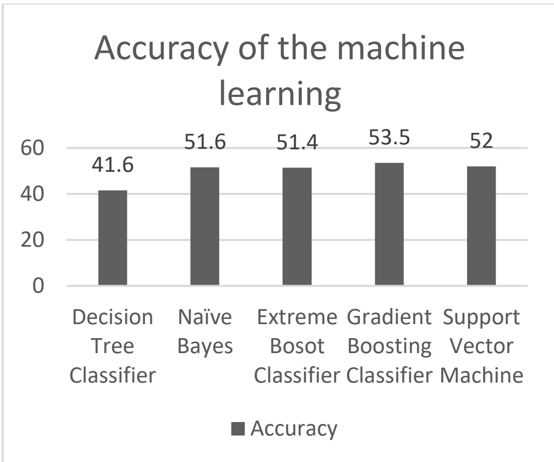

Figure 3: Accuracy of the machine learning algorithm

D. Evaluation and Model Selection

The classification models listed below were tested with the model

- Decision Tree Classifier
- Naïve Bayes
- Extreme Gradient Boost Classifier
- Gradient Boosting Classifier
- Support Vector Machines

The models were chosen based on their performance on three main metrics; accuracy, precision, and recall. **Accuracy** defines the percentage of predictions made by the model that is the same as the actual result. **Precision** represents the percentage of correct predictions the model makes of a certain outcome out of the total number of correct predictions the model makes in total. **Recall** is the total number of accurate predictions the model makes of one type of result concerning the total number of positives of that result within the data set.

$$Accuracy = \frac{total\ number\ of\ correct\ predictions}{total\ number\ of\ predicitons} \times 100$$

$$Precsion\ (Draws) = \frac{total\ number\ of\ correct\ draw\ predictions}{total\ number\ of\ correct\ predicitons} \times 100$$

$$Recall\ (Draws) = \frac{total\ number\ of\ correct\ draw\ predictions}{total\ number\ of\ actual\ draw\ results} \times 100$$

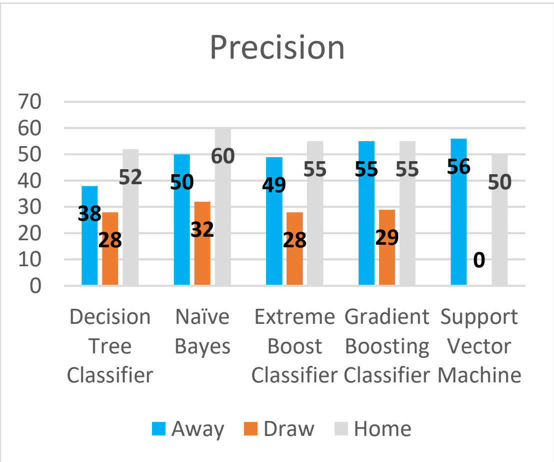

Figure 4: Precision of machine learning algorithms

From the figure 3, it shows that the boosting classifiers; extreme gradient boost classifier, and gradient boosting classifier possessed the highest accuracy of all the models. However, their performance in precision and recall especially in predicting draws was bettered by the other models. Poor precision and recall in predicting draws will mean it would be impossible to trust a draw prediction by this model. Hence, the Naïve Bayes classifier was selected as the best algorithm for the model as it the highest precision and recall for all the possible predictions; home win, away win, or draw (which Naïve Bayes performed particularly well as shown in figures 4 and 5) and excluding the boosting classifiers it had the second-highest accuracy.

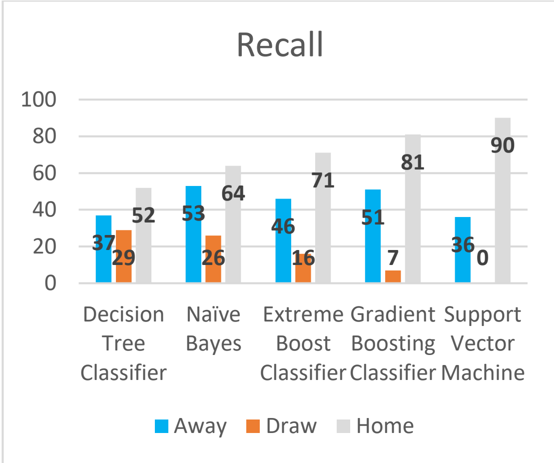

Figure 5: Recall of machine learning algorithms

VI – RESULTS AND DISCUSSION

The final machine learning model selected had an accuracy of 51.6% when it came to predicting the outcome of football games. A test of the model on games that were not yet played at the time resulted in an accuracy of 50% as shown in figure 6. The purpose of the machine learning model is not to encourage gambling but to prevent the gamblers from placing themselves in precarious financial situations. There is a potential controversy of condoning gambling, but it is a small risk to take to protect the health and well-being of people.

Investigating the impact of imagery on addiction, research findings from Block *et al* show that recall of antidrug advertising was associated with a lower probability of marijuana and cocaine use. However, recall of such advertising was not associated with the decision of which drug to use [9]. These findings substantiate that anti-drug advertisements are good for addicts and therefore, it would be useful to incorporate them in the application to help the gamblers.

VII – CONCLUSION AND FUTURE WORKS

In a nutshell, the project is still in progress, but a great deal of progress has been made in the programming of a machine learning model that performs with a 51.6% accuracy. Further work must be performed to implement the psychological aid aspect of the application. With gambling, there are a complex set of harms, co-occurring issues, and population-specific challenges that underscore the need for effective problem gambling (GP) prevention and treatment services [10]. Hence, further work is needed in making questionnaires and interviews to achieve a higher level of qualitative research into the population-specific challenges affecting rural Ghanaians to improve the quality of psychological intervention that is performed on the mobile application. Finally, there is the need to test the application and its response with actual users to gauge their response to find out the success of the Chancer mobile app.

Home Team	Away Team	Result	Model Prediction
Atalanta	Spezia	H	H
Lazio	Juventus	A	A
Fiorentina	AC Milan	H	A
Sassuolo	Cagliari	D	H
Salernitana	Sampdoria	A	H
Bologna	Venezia	A	H
Inter	Napoli	H	H
Genoa	Roma	A	A
Verona	Empoli	H	H
Torino	Udinese	H	D

Figure 6: Serie A results as predicted by Naive Bayes classifier

REFERENCES

[1] Odame, S. K., Quarshie, E. N.-B., Oti-Boadi, M., Andoh-Arthur, J., & Asante, K. O. (2020). Adolescent problem gambling in rural ghana: Prevalence and gender differentiation. *Journal of Gambling Studies*, *37*(1), 83–105. https://doi.org/10.1007/s10899-020-09987-6

[2] Lopez-Gonzalez, H., Estevez, A. and Griffiths, M. (2017). 'Marketing and advertising online sports betting: a problem gambling perspective. *Journal of Sports and Social Issues*, 41(3): 256-272

[3] Glozah, F. N., Tolchard, B., & Pevalin, D. J. (2019). Participation and attitudes Towards gambling in Ghanaian YOUTH: An exploratory analysis of risk and protective factors. *International Journal of Adolescent Medicine and Health*, *33*(4). https://doi.org/10.1515/ijamh-2018-0175

[4] Victorian Casino and Gaming Authority Definition and Incidence of Pathological Gambling Including the Socioeconomic Distribution. Melbourne: Victorian Casino and Gaming Authority; 1997.

[5] St Quinton, T., & Morris, B. (2021). Gambling Prevention Mobile Applications: Understanding the Inclusion and Use of Behaviour Change Techniques. *Journal Of Gambling Issues*, (48), 110-121. https://doi.org/10.4309/jgi.2021.48.5

[6] Markham, F., Young, M. and Doran, B. (2014), 'Gambling expenditure predicts harm: evidence from a venue-level study', *Addiction*, 109(9): 1509–1516

[7] Ofosu, Albert; Kotey, Richard Angelous (2020) : Does Sports Betting Affect Investment Behaviour? Evidence from Ghanaian Sports Betting Participants., Journal of Gambling Issues, ISSN 1910-7595, Centre for Addiction and Mental Health, Toronto, Vol. 43, pp. 61-83,

[8] Constantinou, A., Fenton, N., & Neil, M. (2012). pi-football: A Bayesian network model for forecasting Association Football match outcomes. Knowledge-Based Systems, 36, 322-339. https://doi.org/10.1016/j.knosys.2012.07.008

[9] Block, L., Morwitz, V., Putsis, W., & Sen, S. (2002). Assessing the Impact of Antidrug Advertising on Adolescent Drug Consumption: Results From a Behavioral Economic Model. *American Journal Of Public Health*, *92*(8), 1346-1351. https://doi.org/10.2105/ajph.92.8.1346

[10] Matheson, F., Hamilton-Wright, S., McLuhan, A., Shi, J., Wiese, J., & Kryszajtys, D. et al. (2021). Self-Management Strategies for Problem Gambling in the Context of Poverty and Homelessness. *Journal Of Gambling Issues*, (48), 81-109. https://doi.org/10.4309/jgi.2021.48.4