Using machine learning to predict self-exclusion status in online gamblers on the PlayNow.com platform in British Columbia

Stephanie Finkenwirth*[1], Kent MacDonald*[1], Xiaolei Deng[1], Tilman Lesch[1], Luke Clark[1,2]

*these authors contributed equally to the manuscript

[1]Centre for Gambling Research at UBC, Department of Psychology, University of British Columbia, Vancouver, B.C., Canada

[2]Djavad Mowafaghian Centre for Brain Health, University of British Columbia, Vancouver, B.C., Canada

This is an Author Accepted Manuscript. This paper is not the copy of record and may not exactly replicate the final, authoritative version of the article. Please do not copy or cite without authors' permission. The final article is published as:

Finkenwirth S*, Macdonald K*, Deng X, Lesch T, Clark L. Using machine learning to predict self-exclusion status in online gamblers on the PlayNow.com platform in British Columbia. *International Gambling Studies,* 2020, DOI: 10.1080/14459795.2020.1832132

Corresponding Author

Dr Luke Clark, Department of Psychology, University of British Columbia, 2136 West Mall, Vancouver, B.C., V6T 1Z4, Canada. Tel: 001 604 827 0618. Email: luke.clark@psych.ubc.ca

Funding

This project was funded by a Research Grant from the Province of BC Ministry of Finance (Gaming Policy & Enforcement Branch) awarded to Luke Clark and Tilman Lesch, and by the core funding of the Centre for Gambling Research at UBC.

Disclosures

LC is the Director of the Centre for Gambling Research at UBC, which is supported by funding from the Province of British Columbia and the British Columbia Lottery Corporation (BCLC), a Canadian Crown Corporation. The research receives separate funding from the Province of BC Ministry of Finance (Gaming Policy & Enforcement Branch) to LC and TL. The dataset for this study was provided to the research team by the British Columbia Lottery Corporation under a Non-Disclosure Agreement that prohibits further data sharing. The Province of BC government and the BCLC had no further role in the design, analysis, or interpretation of the study, and impose no constraints on publishing. LC also receives funding from the Natural Sciences and Engineering Research Council (Canada). LC has received a speaker/travel honorarium from the National Association for Gambling Studies (Australia) and the National Center for Responsible Gaming (US), and has received fees for academic services from the National Center for Responsible Gaming (US) and Gambling Research Exchange Ontario (Canada). He has not received any further direct or indirect payments from the gambling industry or groups substantially funded by gambling. He has provided paid consultancy to, and received royalties from, Cambridge Cognition Ltd. relating to neurocognitive testing. TL is now employed by Deloitte Consulting GmbH. KM is now employed by Visier Inc.. SF, KM, XD, TL have no further disclosures to be declared.

LC is a Regional Assistant Editor for International Gambling Studies but was blinded to the review and editorial process for this manuscript.

Data Availability Statement

The dataset was provided in a de-identified format by the British Columbia Lottery Corporation, Data Analytics team in October 2015. Data were given to the Centre (LC and TL) under a Non-Disclosure Agreement that does not allow sharing of the dataset, or reporting of data from individual users. In the manuscript, we provide links to our code, and the exact model values, at github.com/CGR-UBC/PlayNow_VSEprediction_2020.

Acknowledgements

We would like to thank the Social Responsibility and Data Analytics teams at BCLC for their assistance in providing the data.

ABSTRACT

The identification of disordered gambling in the online environment may enable interventions to be targeted to those users experiencing harms. We tested the performance of machine learning in classifying online gamblers with and without a record of voluntary self-exclusion (VSE). We analyzed a one year dataset from PlayNow.com, the provincially-owned online gambling platform in British Columbia, Canada. The primary model compared 2,157 gamblers with a record of VSE enrollment (6 months to 3 years) against 17,526 non-VSE controls, using 20 input variables of gambling behaviour. Machine learning (random forest classifier) achieved an Area Under the Receiver Operating Characteristic curve (AUROC) of 0.75 (SD = 0.01). The input variable with the greatest predictive signal (based on feature importance values) was Variance in Money Bet per Session. Further analyses tested a logistic regression model as a benchmark, and tested the impact of key modelling decisions (including use of a balanced dataset, and data inclusion threshold). Across all models, machine learning was able to predict VSE status with performance between 0.65 – 0.76, using our behavioural inputs. These results provide proof-of-principle data for the applied use of behavioural tracking to identify disordered gambling, and highlight the importance of behavioural inputs reflecting betting variability.

Keywords: online gambling, slot machines, gambling disorder, data science, machine learning, loss chasing

INTRODUCTION

Over the past two decades, online gambling has become a major part of the gambling landscape. Individuals who gamble online typically display elevated levels of disordered gambling (Edgerton, Biegun, & Roberts, 2016; Gainsbury, 2015). Some studies have indicated that this increase is driven by online gamblers engaging in a wider range of gambling forms, including land-based gambling (LaPlante, Nelson, & Gray, 2014; Philander & MacKay, 2014). However, in a recent study, both online-only and mixed-mode (i.e. online and land-based) gamblers displayed higher rates of problem gambling and a range of specific gambling harms, compared to a control group of land-based gamblers (Papineau et al., 2018). The longstanding concerns regarding the risk profile of online gambling arise from a multitude of factors including high availability (especially on mobile devices), card-based payment, and the increased potential to combine gambling with drug or alcohol use (Cotte & Latour, 2009). At the same time, online gambling operators automatically receive large volumes of customer data that is registered against single user accounts. This enables behaviour to be monitored across time (termed ‘behavioural tracking’) and creates an opportunity for online gambling operators to identify and respond to users displaying signs of disordered gambling (Deng, Lesch, & Clark, 2019; Edgerton et al., 2016; Griffiths, 2012). Advances in data science and the application of these techniques to psychology and psychiatry are ideally suited to this objective (Deng et al., 2019), and regulators are increasingly considering mandatory requirements for data analytics as a means of reducing gambling harm. Some concerns have been raised about the transparency of predictive techniques that rely on complex, multivariate ‘black box’ algorithms (Coussement & De Bock, 2013).

Identifying Disordered Gambling from Behavioural Tracking Data

Prior research on behavioural tracking of online gambling data has relied on a small number of datasets, chiefly from European websites. A series of studies by the Harvard Division of Addictions have used data from *bwin.party*, a European website that specializes in sports betting but also offers online casino games. One *bwin* dataset comprises users who opened accounts on *bwin* in February 2005 (n = 48,114) and were followed over a 2 year period (Labrie, Kaplan, Laplante, Nelson, & Shaffer, 2008; Nelson et al., 2008; Shaffer, Peller, LaPlante, Nelson, & LaBrie, 2010; Xuan & Shaffer, 2009). A key challenge in working with online gambling data is the verification of which users actually have gambling problems. Although some studies have approached this directly by hosting a problem gambling screening instrument on the gambling platform (Excell et al., 2014; LaPlante et al., 2014; Tom, LaPlante, & Shaffer, 2014), response rates tend to be low. For example, only 2% of 100,000 subscribers to *bwin* voluntarily completed the Brief Biosocial Gambling Screen in Tom, et al. (2014), raising concerns about the representativeness of an algorithm built on such individuals. An alternative marker is account closure, where it may be further possible to compare users based on stated reasons for account closure (e.g. gambling problems versus dissatisfaction with the platform) (Braverman & Shaffer, 2012; Luquiens et al., 2018; Philander, 2014). Other possible 'red flags' include complaints to the website or requests to increase spending limits (Braverman, LaPlante, Nelson, & Shaffer, 2013; Gray, LaPlante, & Shaffer, 2012).

The present report uses voluntary self-exclusion (VSE) as an indicator of gambling harm. VSE programs enable gamblers to ban themselves from a gambling facility or operator for a set period of time (in the British Columbia program, the gambler selects a period between 6 months and 3 years duration). Past research indicates that the majority of gamblers enrolled in VSE programs meet criteria for problem gambling (Ladouceur, Jacques, Giroux, Ferland, & Leblond, 2000; Motka et al., 2018), supporting the validity of VSE status as an indicator of disordered gambling for use in predictive modelling. For example, in an evaluation of 269 gamblers enrolled in the self-exclusion program in British Columbia, 74% scored in the problem gambling category on the Problem Gambling Severity Index (McCormick, Cohen, & Davies, 2018). Although research on online VSE is more limited, of 259 VSE gamblers on a European website, 76% met criteria for problem gambling (Hayer & Meyer, 2011). Nevertheless, these figures serve to highlight that a proportion of VSE enrollees do not evidence problem gambling on screening questionnaires. In addition, only a small minority of problem gamblers appear to enroll in VSE programs, for example compared to the prevalence of gambling problems among casino patrons (Dragicevic, Percy, Kudic, & Parke, 2015). Implementation of predictive modelling by gambling operators will likely require convergent evidence across multiple indicators of gambling harm.

Behavioural Predictors and Machine Learning

Online gambling datasets often contain millions of individual bets. In extracting quantitative behavioural markers for each user, bets are usually aggregated by day (e.g. number of active betting days) and/or by session (e.g. average bets per session). Sessions are usually defined as periods of uninterrupted play, where each bet follows the previous bet within a predefined period of time, e.g. 30 minutes. Early *bwin* studies established that disordered gambling is associated with higher gambling *frequency* (e.g. active betting days) and gambling *intensity* (e.g. average bet size) (Labrie & Shaffer, 2011; Nelson et al., 2008). Measures of *variability* are also identified as an important predictor; for example, high-risk users tend to vary their bet size more (i.e. the standard deviation of the wager size) (Braverman et al., 2013; Braverman & Shaffer, 2012). A distinct form of variability (sometimes termed breadth or diversity) is the number of gambling products that the user engages with (Adami et al., 2013; LaPlante et al., 2014). Some studies have also sought to characterize gambling trajectories; for example, increasing bet size and increasing losses were observed in the days leading up to account closure (Xuan & Shaffer, 2009).

Several studies have entered multiple behavioural predictors in multivariate models (Braverman et al., 2013; Gray et al., 2012; Luquiens et al., 2018). In a *bwin* analysis of 2,042 individuals with various problem gambling indicators (and 2,014 controls), Gray et al. (2012) applied discriminant function analysis on 27 input variables. They found that non-monetary variables reflecting frequency and intensity (active betting days, bets per day) best distinguished the cases with a variety of ‘red flags’ for problem gambling (including requests for account closure, account re-opening, or changes to deposit limits), particularly on variables pertaining to live action sports. In a distinct European dataset from the operator GTECH, users who later self-excluded from the program displayed higher gambling losses and greater variability in losses, but did not differ on bets per day or the number of different games played (Dragicevic et al., 2015). Thus, it remains unclear which behavioural variables are robustly associated with markers of problem gambling, and additional online gambling datasets are needed to establish generalizability.

Machine learning techniques can be used for building predictive models, and have garnered much interest in psychology and psychiatry (e.g. Ahn and Vassileva, 2016; Chekroud et al., 2016). In cases where the outcome for each participant is known prior to analysis, the modelling process is termed ‘supervised learning’: the algorithm is presented with input data along with the outcome categorization, and then ‘learns’ which patterns in the input data are associated with the outcomes. This is formulated as a set of mathematical rules (e.g. in the form of decision trees), which can in turn be used for predicting outcomes in new input data (see Dwyer et al., 2018 for review). Advantages of machine learning methods over linear models such as logistic regression is that machine learning algorithms are designed for complex multivariate data and include interactions by default, rely on fewer assumptions regarding the data (e.g. normality, linear relationships), and have built-in procedures (i.e. cross-validation) to increase generalizability (Dwyer et al., 2018; Gowin et al., 2019).

Previous studies testing such procedures in online gambling datasets have found mixed success. Using a publicly available *bwin* dataset, Philander (2014) trained the model on the comparison of 176 account closers who cited gambling problems and 354 account closures who gave other reasons. The input variables comprised the behavioural markers of frequency, intensity, variability, and trajectory from Braverman & Shaffer (2010), plus some demographic variables. Philander trained the model on 70% of the dataset and tested the resulting algorithm on the remaining 30%. Area Under the Receiver Operating Characteristic curve (AUROC) values were reported as a performance metric. The Receiver Operating Characteristic (ROC) curve plots the true positive rate (sensitivity) against the false positive rate (i.e. 1 - specificity), across all possible discrimination thresholds. An AUROC of 1 indicates perfect prediction, while a random classifier would obtain an AUROC of 0.5. In the Philander study, across 9 different algorithms that included a random forest classifier, AUROC values varied from 0.50 - 0.55, offering only a modest benefit over chance. It is possible that this poor performance stemmed from using controls who were also account closers. A recent study of online poker players found that the stated reason for self-exclusion (gambling problems vs commercial reasons) yielded similarly low classification performance based on behavioral inputs, across three machine learning models (highest AUROC 0.57) (Luquiens et al., 2018).

A later study using a dataset from GTECH (Percy et al., 2016) employed self-exclusion as the indicator for problem gambling, training a model to discriminate 176 VSE gamblers against 669 controls. As inputs, 33 variables reflecting frequency, intensity, variability, trajectory, as well as basic demographics were entered. A random forest classifier had an AUROC of 0.79, which is considerably higher than performance in Philander (2014) and Luquiens et al (2018). Similar analyses using other sources of input data have also found reasonable levels of classificatory performance, based on payment-related data (e.g. deposits and withdrawals from the online account) (AUROC 0.72) (Haeusler, 2016) and text analysis of email correspondence with the operator (AUROC 0.78) (Haefeli, Lischer, & Haeusler, 2015). Although promising, the Percy et al. analysis provided no information on the relative importance of their behavioural input variables. Information on 'feature importance' values would help to allay concerns regarding the lack of transparency of machine learning as a 'black box' procedure (Coussement & De Bock, 2013), and will also help to resolve disparities in model performance across different datasets.

In the current study, we apply machine learning to online gambling data from PlayNow.com, the provincially-owned online gambling platform in British Columbia, Canada. As input data, we selected 20 behavioral variables aggregated by day or session, reflecting frequency, intensity, and variability of gambling. In comparison to prior work in which some behavioural variables required pre-processing (e.g. definition of baseline periods and significance testing of trajectories, Percy et al., 2016), we deliberately selected coarse inputs that required minimal pre-processing. Our target group was users with a record of ever having enrolled in the provincial VSE program (these records could reflect prior enrolment that had expired, or enrolment within the 1 year data window). In our primary model, we trained a random forest classifier on a large, unbalanced dataset of 2,157 VSE gamblers against 17,526 non-VSE controls. Our first objective was to establish the overall predictive performance (AUROC) of the model based on behavioural inputs, and the second objective was to establish which individual behavioural variables were most predictive of VSE status, based on feature importance values. We report a series of sensitivity analyses to test the impact of key decisions in the modelling process.

METHODS

Our dataset contains one year (1 October 2014 to 30 September 2015) of online gambling activity from BC residents using the eCasino section of PlayNow.com platform. The PlayNow.com platform was introduced in BC by the British Columbia Lottery Corporation (BCLC) in 2004 and was the first provincially-operated, regulated gambling website in North America. In 2015, there were an estimated 265,000 registered users on PlayNow.com in BC (Province of BC, 2015). The eCasino section of the platform contains online slot machine and table games. In a previous analysis of 1 month of data from PlayNow.com, 97% of bets on the overall platform were placed in the eCasino section (Clark & Lesch, 2018).

The dataset was provided in a de-identified format by the BCLC Data Analytics team in October 2015 following a request by the Centre for Gambling Research to the BCLC Social

Responsibility team. Data were given to the Centre under a Non-Disclosure Agreement that does not allow sharing or reporting of data from individual users. The Behavioural Research Ethics Board at the University of British Columbia approved the protocol for the secondary analyses. The user agreement upon registering with PlayNow.com states that gambling data may be shared with third parties for research purposes. In preparing the data for transfer, the BCLC Data Analytics team assigned each user a unique, randomly generated ID number that was linked to their betting behaviour. The time and date of all bets was recorded. The dataset was stored on a secure server (Westgrid) hosted in Canada.

The full dataset comprises 30,902 individual customers, placing over half a billion (575,470,087) individual bets. In the dataset, 2,458 users had a record of self-exclusion. The BC self-exclusion program was established in 1998, and in 2017, there were 9,565 individuals enrolled in program (https://www2.gov.bc.ca/gov/content/sports-culture/gambling-fundraising/news-updates/2018-02-14). On enrolling in the BC self-exclusion program, gamblers select a duration of 6 months, 1 year, 2 years or 3 years. Shorter term bans were not available at the time of our study (although a 1-14 day 'lock-out' function was introduced subsequently). Gamblers can enroll in VSE either at land-based gambling venues, or – as in the case of our dataset - through the PlayNow.com website. We note that we did not receive data on the length of the self-exclusion period or the number of prior self-exclusions, and our self-exclusion variable could reflect either a *prior* VSE enrolment that had expired, or a user who enrolled in VSE at some point during our data window. The data inclusion thresholds (see below) ensured that participants were active users on the PlayNow.com platform within the 1 year period under study.

Preprocessing and Performance Metrics

Some users had negligible levels of activity on the platform during the 1 year window. For inclusion, we set a minimum data threshold of at least 200 bets. This resulted in 2,157 self-excluders and 17,526 controls in our primary model. We recognize that this threshold is arbitrary, and it is an empirical question how much data per user is required to build meaningful models and predict indicators of disordered gambling. As the eCasino comprises online slot machine and table games, 200 bets could occur in less than 30 minutes of gambling based on typical event frequencies, and thus within a single session. As we discuss further below, we tested the impact of thresholding in a sensitivity analysis using an alternative threshold of at least 10 sessions (see also Percy et al. 2016 who used this alternate threshold in their study).

As inputs, we identified 20 betting variables, reflecting multiple aspects of gambling frequency, intensity, and variability (see Table 1 for variable definitions and Supplementary Table 1 for descriptive statistics), as guided by previous research on online gambling behavioural markers (Braverman et al., 2013; Dragicevic et al., 2015; Gray et al., 2012; Luquiens et al., 2018; Percy, França, Dragičević, & d'Avila Garcez, 2016). These were derived by aggregating bet-by-bet data across days or sessions, and in addition by calculating the variance of some indicators (e.g. Variance in Bets per Session). For the scope of this paper, we do not conduct extensive feature engineering; for example, we do not establish baseline windows to quantify trajectories (c.f. Percy et al., 2016), or seek to derive/infer psychological constructs such as loss chasing (Adami et al., 2013). Whether the incorporation of such variables improves model performance is an important question for future research.

Table 1: Input variables of gambling behaviour included in the primary model. Italicized variable names reflect the 6 variables identified by the variable selection algorithm.

Variable Name	Description
Days Gambled	Total number of days in which at least one bet was made by the same customer
Total Sessions	Total number of sessions. A session describes a sequence of bets made by the same customer, where each bet is placed no more than 30 minutes after the previous bet
Sessions per Day	*Total Number of Sessions divided by Days Gambled*
Total Bets	The total number of bets made by the same customer
Bets per Day	*Total Bets divided by Days Gambled*
Bets per Session	Total Bets divided by Total Sessions
Variance in Bets per Session	Variance of Bets per Session
Distinct Games per Session	The number of unique games (within the eCasino section) played by the same customer divided by Total Sessions
Variance in Distinct Games per Session	Variance of Distinct Games Per Session
Total Money Bet	Total amount of money wagered by the same customer in Canadian dollars
Money Bet per Session	Total Money Bet divided by Total Sessions

Variance in Money Bet per Session	*Variance of Money Bet per Session*
Average Session Length	Mean session length (in seconds)
Variance in Average Session Length	Variance of Average Session Length
Total Money Bet from Promotional Offers	Total amount of "Promotional" money wagered by the same customer in Canadian dollars
Money Bet from Promotional Offers per Session	*Total Money Bet from Promotional Offers divided by Total Sessions*
Total Net Loss	The total amount of money lost by the same customer in Canadian dollars
Net Loss per Session	Total Net Loss divided by Total Sessions
Total Net Win	*The total amount of money won by the same customer in Canadian dollars*
Net Win per Session	Total Net Win divided by Total Sessions

In line with previous studies, we report the AUROC as our primary performance metric. AUROC is widely used to evaluate the quality of probabilistic predictions, as it is independent of any particular classification threshold, and random performance remains an AUROC = 0.5 in unbalanced datasets. The advantage of probabilistic predictions compared to discrete class predictions is that the former offers information on the certainty of each prediction. That is, the class of interest is predicted with a degree of certainty that ranges between 0 and 1. This information is lost when considering classification-based metrics (such as sensitivity and specificity), because predicted probabilities need to be projected onto a discrete (in our case binary) scale first. However, to aid comparison with previous work, we report sensitivity and specificity as secondary performance metrics in Tables 2 and 3 in addition to AUROC. [1] It is worth emphasizing that AUROC accounts for both the true positive rate as well as the true negative rate (at each threshold), which can lead to overly optimistic evaluation of an algorithm in cases where correctly classifying negatives is less important than correctly classifying positives (Davis & Goadrich, 2006). For our objectives, it is not clear whether it would be worse to mistakenly label a non-VSE gambler as a VSE target, or vice versa. We note that alternatives to AUROC also exist for extremely unbalanced datasets, such as the precision-recall curve (Sun, Wong, & Kamel, 2009), although we have opted to test the impact of unbalancing more directly, by running a sensitivity test with a balanced dataset, as outlined below.

Analysis Plan

Models were run using the python library scikit-learn (scikit-learn.org). Github code and exact model values are available at github.com/CGR-UBC/PlayNow_VSEprediction_2020. Our primary model was a random forest classifier on the full set of behavioural input variables (see Table 1 and Supplementary Table 1), based on 2,157 self-excluders and 17,526 controls. We

[1] For the purpose of classification, we use an optimized cutoff value to account for the skew in our base rates in the unbalanced case. The optimized cutoff value is found by maximizing the difference between True Positive Rate and False Positive Rate, the two axes on the ROC curve on the training dataset.

focus on these unbalanced groups in order to make full use of the available data, acknowledging that self-exclusion, and disordered gambling more broadly, are relatively rare events, and thus real-world data for predictive modelling are likely to be unbalanced. A random forest classifier is an ensemble machine learning method based on independently built decision trees. Ensemble methods combine several base learners to achieve higher generalizability compared to a single model. The random forest algorithm arrives at a final result by averaging the predictions of all decision trees in the ensemble. We ran a logistic regression on the same data as a benchmark using classical statistics that assume linearity.

Model performance estimation requires testing on unseen data, and with careful attention to avoid information leakage (Dwyer et al., 2018). This can be achieved with k-fold cross-validation (Hastie, Tibshirani, & Friedman, 2009), a procedure in which the dataset is initially divided into k folds of approximately the same size, following a specified sub-sampling regime (e.g. with or without stratification). Typical values for k are 5 or 10. The training and testing process is repeated k times, such that all k folds will have been part of the test set once. A predefined performance metric, AUROC in our case, is calculated for each repeat. This results in k performance values, which are then averaged to attain a single value to evaluate overall model performance. In cases where the training procedure involves optimization steps, such as the search for optimal hyperparameters, *nested* cross-validation should be used (see e.g. Dwyer et al., 2018) to obtain unbiased performance estimates. It consists of two cross-validation loops, nested into each other: the outer loop is used for testing, while the complete training procedure (including hyperparameter optimization) takes place in the inner loop. In our study, both the random forest classifier and logistic regression have hyperparameters that can be optimized. Thus, we employed stratified 10-fold nested cross validation for both models.

As mentioned above, we also report classification-based performance metrics, i.e. sensitivity (true positive rate) and specificity (true negative rate) in Tables 2 and 3. To calculate these measures we obtain predicted probabilities for the entire dataset by using the best performing model specifications from the nested cross-validation procedure within a simple 10-fold loop, in turn training and predicting on each of the ten folds. For each predicted fold we find the optimal classification threshold by maximizing the difference between true positive rate and false positive rate. To avoid overfitting, we average the resulting cutoff values over all folds before using it to classify the entire dataset.

For the random forest classifier, feature importance rankings were used to establish which input variables were most influential in predicting self-exclusion. Feature importance values are defined as the total decrease in node impurity for a given tree, averaged over all trees of the random forest ensemble (Breiman, Friedman, Stone, & Olshen, 1984). However, some of our input variables were moderately to highly inter-correlated, as measured by Pearson's correlation coefficients (see Supplementary Figure 1), for example Total Sessions and Days Gambled. Although random forest classification performance is not affected by collinearity of inputs, feature importance values would be affected. Thus, for calculating feature importance values, we first used a variable selection algorithm to identify a reduced number of inputs that were sufficiently uncorrelated with each other. Feature selection based on correlation coefficients typically relies on dropping all but one of a set of correlated features. There are several methods to decide on which variables to drop, including correlation with target variable and random choice. We opted for developing a custom algorithm that did not rely on utilizing the target variable and could be easily integrated into our machine learning pipeline (see our Python code in Github repository /_feature_selection_class.ipynb). The program iteratively searches for and eliminates subsets of variables around 'nucleus' variables, where the correlation coefficient between a nucleus variable and each variable in its surrounding subset is higher than some constant (0.3 in our case). Retaining the nucleus variables preserves most of the information captured (since, by definition, the surrounding variables are all highly correlated with the nucleus), while reducing model complexity. The procedure was integrated into the training pipeline executed in the inner cross-validation loop, as outlined above.

In summary, our primary model was a random forest classifier on the 20 input variables, using the unbalanced dataset of 2,157 self-excluders and 17,526 controls, with the 200-bet data inclusion threshold. We report a series of sensitivity analyses to test the impact of key decisions in our primary model. First, we re-ran the random forest classifier on the reduced number of inputs identified by the variable selection algorithm. Second, we tested the impact of unbalanced data by creating a fully balanced dataset; we randomly selected a subset of 2,157 controls. We use the 'choices' method of the 'random' library in Python to select a subset of the control of the same size as the target group (for our open code, see Github link). We note that this 'down-selection' inherently loses much data, and that some more sophisticated methods exist for handling unbalanced sets, including Synthetic Minority Oversampling Technique, SMOTE (Percy et al., 2016). Third, we tested whether our results depended on the 200-bet data inclusion threshold. Conceivably, thresholding data based on the number of bets could bias the model towards particular input variables (e.g. those related to bet intensity). Our sensitivity check used an alternative data inclusion threshold of at least 10 sessions, based on Percy et al., (2016). A session was defined as a period of gambling from the time of user login, until either the user logged out, or automatic log-out occurred after 30 minutes of inactivity. This threshold identified 1,776 self-excluders and 13,470 controls. Fourth, we ran a logistic regression as a classical linear benchmark to explore the advantages of random forest as a form of machine learning.

RESULTS

Results for the random forest classifier are shown in Table 2 and for the logistic regression benchmark model in Table 3. In the primary model (i.e. the unbalanced dataset, 20 input variables, 200-bet threshold), the random forest classifier achieved an AUROC of 0.75 (SD = 0.01). Under the same conditions, logistic regression performed poorly, with an AUROC of 0.39 (SD = 0.03).

Table 2: Random Forest Classification modeling results

	Unbalanced data		**Balanced data**	
	2,157 vs. 17,526	**1,776 vs. 13,470**	**2,157 vs. 2,157**	**1,776 vs. 1,776**
	>200 bets	**>10 sessions**	**>200 bets**	**>10 sessions**
All variables (20)				
AUROC	0.75 (0.01)	0.73 (0.02)	0.76 (0.02)	0.75 (0.02)
sensitivity	0.69	0.67	0.73	0.74
specificity	0.65	0.69	0.63	0.65
Selected variables (6)				
AUROC	0.70 (0.01)	0.65 (0.01)	0.72 (0.03)	0.69 (0.02)
sensitivity	0.63	0.60	0.63	0.65
specificity	0.59	0.70	0.63	0.69

AUROC = Area under the Receiver Operating Characteristic curve

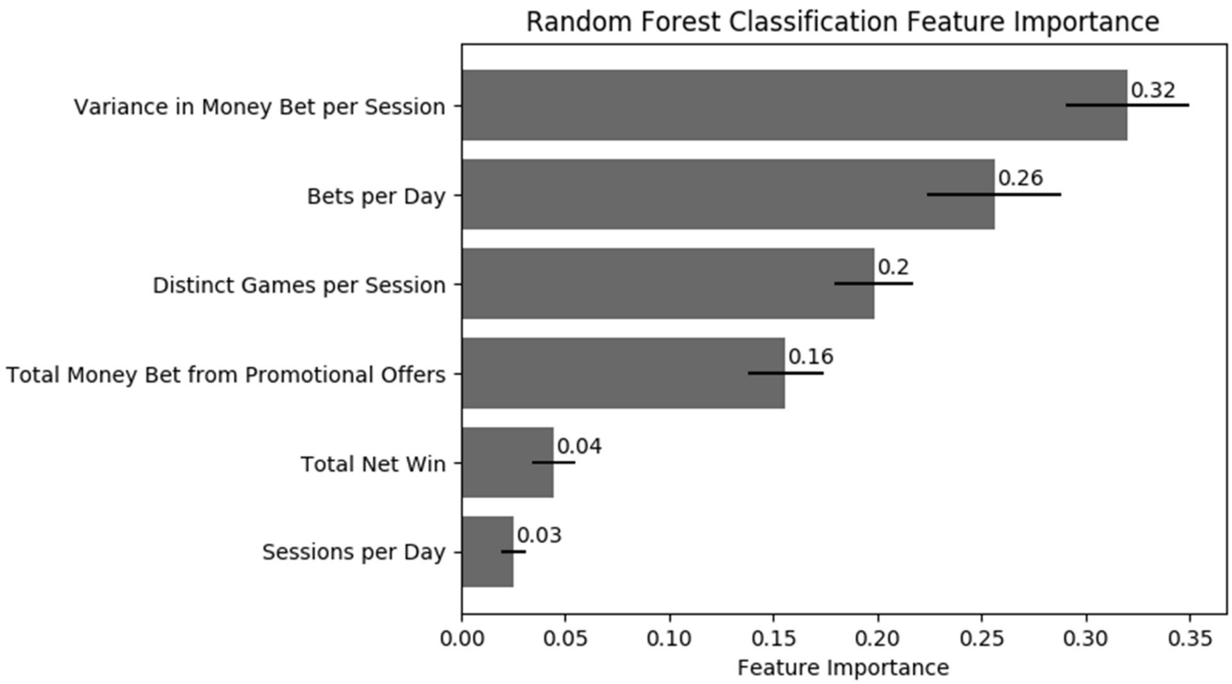
Table 3: Logistic Regression modeling results

	Unbalanced data		**Balanced data**	
	2,157 vs. 17,526	**1,776 vs. 13,470**	**2,157 vs. 2,157**	**1,776 vs. 1,776**
	>200 bets	**>10 sessions**	**>200 bets**	**>10 sessions**
All variables (20)				
AUROC	0.39 (0.03)	0.41 (0.03)	0.68 (0.02)	0.65 (0.02)
sensitivity	0.10	0.03	0.61	0.60
specificity	0.81	0.96	0.66	0.67
Selected variables (6)				
AUROC	0.69 (0.02)	0.63 (0.07)	0.70 (0.01)	0.68 (0.02)
sensitivity	0.46	0.41	0.67	0.68
specificity	0.71	0.76	0.61	0.56

AUROC = Area under the Receiver Operating Characteristic curve

The variable selection algorithm identified 6 variables that were sufficiently uncorrelated (i.e. had correlation coefficients lower than 0.3). These variables are italicized in Table 1. Figure 1 shows the feature importance ranking from the random forest classifier on the reduced variable set. The Variance in Money Bet per Session had the highest feature importance, accounting for 32% of the predictive signal. The feature importance rankings for the full set of input variables is shown in Supplementary Figure 2, in which Money Bet per Session and Variance in Money Bet per Session each accounted for 8% of signal.

Figure 1: Feature importance values from the random forest classifier, based on the reduced set of 6 input variables identified by the variable selection algorithm



For the machine learning models in Table 2, the variable selection algorithm had a slight negative effect on classification performance (e.g. AUROC 0.75 (SD = 0.01) in the primary model to 0.70 (SD = 0.01) based on the 6 inputs). This detrimental effect was greatest for the random forest classifier with the alternative 10 session data threshold, in which the AUROC decreased from 0.73 (SD = 0.02) in the full variable set to 0.65 (SD = 0.01) in the reduced variable set. Conversely, logistic regression performance increased substantially on the reduced variable set, with the AUROC rising to 0.69 (SD = 0.02) (see Table 3).

The balanced dataset was associated with slightly higher performance. In the primary model, performance increased from 0.75 (SD = 0.01) to 0.76 (SD = 0.02) on the balanced data. As this falls within the SD range, the difference may represent chance. Lastly, the alternative 10 session threshold performed slightly worse than the primary model, with AUROC ranging from 0.65 (SD = 0.01) to 0.73 (SD = 0.02).

DISCUSSION

This paper tested the predictive performance of machine learning in classifying self-exclusion status in online gamblers. As inputs, we used behavioral variables derived from betting data aggregated by day or by session. In our primary model, machine learning using a random forest classifier achieved an AUROC of 0.75. Given the coarse nature of our behavioural inputs, we consider this to be promising 'proof of principle' data for the development of applied algorithms. We used feature importance values (based on a smaller set of inputs that were uncorrelated with one another) to investigate which variables were most predictive of VSE status. Variance in Money Bet per Session had the highest feature importance, accounting for 32% of the predictive signal. We performed a number of sensitivity checks of our primary model, to test the impact of key modelling decisions including the use of the variable selection algorithm, using a balanced dataset, and changing the data inclusion threshold. Across these models, machine learning performance varied from 0.65 to 0.76. We infer that model performance is somewhat sensitive to these choices, which should be taken into consideration when building predictive models of gambling risk.

We tested a logistic regression model as a linear benchmark against which the benefits of a machine learning approach could be compared. The logistic regression performed particularly poorly under equivalent conditions to the primary model, with an AUROC of 0.39. The performance of the logistic regression model improved substantially when tested on the smaller set of inputs identified by the variable selection algorithm (AUROC = 0.69) and when tested on the balanced dataset (AUROC = 0.68). By contrast, the performance of the machine learning model decreased slightly on the smaller set of inputs (AUROC = 0.70), so that with the smaller set of inputs and balanced dataset, the logistic regression and random forest models performed almost identically. It is likely that the logistic regression performed poorly in the main model because of high multicollinearity in the full set of inputs, along with other violations of assumptions of traditional models, including normality of distributions and linearity of predictor relationships (see also Gowin et al. 2019). It is also likely that the performance of the logistic regression could be improved with further modelling, and in this sense, our comparison somewhat disadvantages the benchmark model. Nevertheless, by our interpretation, it is a strength of machine learning that it is largely robust to these statistical factors, and this is clearly beneficial for its application to real-world data (i.e. unbalanced groups when predicting rare events, based on often collinear inputs).

Two previous studies testing machine learning procedures for gambling risk prediction yielded mixed results. On the GTECH dataset, Percy et al. (2016) reported an AUROC of 0.79. Although our primary model achieved slightly lower performance than this, Percy et al. (2016) included basic demographic predictors (age and gender) as inputs that were not available in our dataset, as well as some additional behavioural inputs that represented gambling trajectories relative to a baseline period. Our performance exceeds that reported by Philander (2014) on the *bwin* dataset, for which AUROC varied from 0.50 - 0.55, representing only modest improvements above chance performance. Philander's analysis on the *bwin* dataset focused on account closers, with those citing gambling problems as their reason for closure identified as the target group. The superior performance seen here supports the interpretation by Luquiens et al (2018), that gamblers' stated reasons for account closure may offer little discriminatory power. For example, gamblers who cite dissatisfaction with the website as their reason for closure may still have elevated levels of gambling problems. It also remains possible that idiosyncratic features of individual datasets and the operationalizing of specific variables could contribute to these performance differences. These issues can only be resolved through increased research access to other online gambling datasets. We note that online gambling operators are likely to have access to a far wider array of information than was included here, and which would likely strengthen model predictions in the future. These include demographic data, analysis of banking transactions into the online gambling account (e.g. multiple deposits within a short interval, or failed deposits) (Haeusler, 2016), engagement with responsible gambling tools on the platform (e.g. adjustments to limit settings), or analysis of correspondence between the gambler and the operator (Haefeli et al., 2015).

Machine learning is sometimes criticized for its 'black box' nature; how could a gambling operator explain a high risk rating to a user? Our analyses address this question of transparency by reporting feature importance values, in order to characterize which input variables are most important in predicting VSE status. It is striking that the Variance in Money Bet per Session accounted for over 32% of the predictive signal (using the smaller set of inputs that were sufficiently uncorrelated). Notably, this is an expenditure-related variable, and gambling harms typically arise from financial losses and their repercussions (Blaszczynski & Nower, 2002; Langham et al., 2016). What may be more unexpected is that it is the *variability* in expenditure that is predictive in our data. This may reflect loss chasing and/or binging tendencies in which the *episodic* breakdown of control may be a key marker of disordered gambling. Bet variability has also been associated with alcohol consumption (Leino et al., 2017), which is challenging to detect in the online environment. It is also possible that this fluctuating bet pattern may reflect inconsistent access to gambling funds, e.g. following paydays. Adami et al. (2013) proposed a ‘sawtooth’ pattern of risky gambling, characterized as gradual bet escalation followed by a collapse as funds are exhausted. These more specific profiles, based on a detailed, within-session behavioural analysis, could be built into subsequent models. Given the ability of random forest models to leverage non-linear features of the data, we also note that the relationship between bet variability and VSE status may be complex, and emerging techniques allow for more detailed characterization of feature importance relationships (Garcia, Dragičević, Percy, & Sarkar, 2019; Lundberg & Lee, 2017).

The likely application of machine learning models is to create risk scores for online gamblers. These scores may enable gambling operators in identify high-risk gamblers *without* a record of self-exclusion, based on their statistical resemblance to self-excluders as the target group. In an analysis of the Swedish PlayScan tool (Wood & Wohl, 2015), 779 online gamblers were informed about their risk level based on an undisclosed and proprietary ‘traffic light’ algorithm. Users receiving the intermediate risk rating significantly reduced their wagering over a 24-week follow-up, although behaviour was not significantly reduced by the feedback in the highest risk group. Nevertheless, the derivation of risk categories involves a number of complex, and potentially arbitrary, decisions. Supplementary Figure 3 shows the histogram of predicted probabilities of self-exclusion (i.e. 'risk score') from our primary model. Inspection of this histogram does not reveal any obvious clusters or kinks for separating a high-risk group. Thus, in our view, the proposed thresholds for different risk categories would require further empirical validation.

Our findings should be interpreted in relation to a number of limitations. First, our analyses were based upon 1 year of data from online slot machine and table games, contained within a wider provincial platform. Behavioural predictors are likely to vary across game types (Brosowski, Meyer, & Hayer, 2012) and our analyses were limited in quantifying breadth of gambling involvement only in the eCasino section (Adami et al., 2013; LaPlante et al., 2014). Online gamblers may differ from land-based gamblers in their psychological profile (Blaszczynski, Russell, Gainsbury, & Hing, 2016; Papineau et al., 2018), and online gamblers who use a regulated platform such as PlayNow may differ from those who opt to use unregulated ('offshore') websites (Gainsbury, Abarbanel, & Blaszczynski, 2019). Certainly, as an analysis of behaviour from a single website (and subset of gambling products), these data may offer only a snapshot of the wider behaviour of these gamblers. Second, our self-exclusion indicator was both retrospective and heterogeneous; these were cases with any record of self-exclusion prior to, or during, the one year period, and selected VSE periods could vary from 6 months to 3 years. Future analyses may assess prospective prediction of self-exclusion, and rates of re-enrolment in VSE after expiry (Luquiens et al., 2018). For future datasets in which VSE status could be predicted prospectively, chronological hold-out testing offers an alternative approach to cross-validation, in which the model is trained on the older data and tested on the most recent data. Third, our input variables did not include demographic variables, and it is likely that younger age and male gender may be over-represented (to an unknown extent) in our target group (Dragicevic et al., 2015; Hayer & Meyer, 2011). With regard to our feature importance analysis, our approach could favor variables that are measured on a larger scale or range of discrete values (Strobl, Boulesteix, Zeileis, & Hothorn, 2007), as was the case here for Variance in Money Bet per Session. Nevertheless, our findings indicate that variability metrics are important to include in the set of input variables for predicting markers of disordered gambling.

A final point concerns the limitations of self-exclusion. Narrowly, our algorithm was built to predict self-exclusion status, and it is an empirical question whether such an algorithm has utility in predicting disordered gambling more broadly. In support of VSE as a proxy marker for disordered gambling, most people who self-exclude are seen to have gambling problems on screening instruments, including in the jurisdiction where this research was conducted (74% in McCormick et al. 2018). In online poker gamblers, the gambling spend in the month leading up to self-exclusion was actually higher than the average monthly spend in group with known problem gambling (Luquiens et al., 2018). At the same time, some self-excluders do not display evident gambling problems, and a significant proportion of problem gamblers do not self-exclude, raising the possibility that self-excluders may represent a specific subtype of problem gambler (Dragicevic et al., 2015; Motka et al., 2018). Since 2015, the BCLC's platform introduced a 1-14 day 'lock-out' feature as a low-barrier alternative to VSE. It is an empirical question to what extent lock-out also serves as a (distinct) marker of gambling risk from VSE. As different methodological issues constrain most proxy markers (e.g. account closure, or sample bias when hosting screening tools on gambling platforms), we propose that convergent data should be sought across multiple markers in order to 'triangulate' the construct of disordered gambling.

This research has focused on the identification of disordered gambling in the online environment, seen through the lens of predicting self-exclusion. We acknowledge that identification is distinct from the nature or effectiveness of any *intervention* that might be built on such identified risk status. Possible interventions include the communication of at-risk status (Wood & Wohl, 2015), feedback on recent expenditure (Auer & Griffiths, 2016; Wohl, Davis, & Hollingshead, 2017), or pop-ups indicating time on device (Auer, Malischnig, & Griffiths, 2014). In our view, our observations that machine learning can classify self-exclusion status using relatively coarse behavioural inputs with performance up to AUROC 0.76, establishes proof-of-principle data for the applied use of behavioural tracking to identify disordered gambling. In addition to the recent attention on replication in gambling research as an important exercise in itself (Wohl, Tabri, & Zelenski, 2019), our findings extend prior research on predictive algorithms (e.g. Percy et al., 2016; Philander, 2014) in at least three ways, by reporting i) on a larger sample of self-excluders, ii) one of the first datasets from North American jurisdiction, iii) feature importance rankings of behavioural inputs, along with testing the impact of a number of key modelling decisions. Behavioural inputs reflecting variability appear to be powerful and should be incorporated in future research. Further signals for risk prediction may be detectable in a more fine-grained analysis of bet-by-bet behaviour, as well as the further tiers of information that are available to online gambling operators. Using such information, future refinement of these algorithms may be capable of predicting risk of disordered gambling with a high degree of accuracy.

References

Adami, N., Benini, S., Boschetti, A., Canini, L., Maione, F., & Temporin, M. (2013). Markers of unsustainable gambling for early detection of at-risk online gamblers. *International Gambling Studies*, *13*(2), 188–204. https://doi.org/10.1080/14459795.2012.754919

Ahn, W. Y., & Vassileva, J. (2016). Machine-learning identifies substance-specific behavioral markers for opiate and stimulant dependence. *Drug and Alcohol Dependence*, *161*, 247–257. https://doi.org/10.1016/j.drugalcdep.2016.02.008

Auer, M., & Griffiths, M. D. (2016). Personalized behavioral feedback for online gamblers: A real world empirical study. *Frontiers in Psychology*, *7*(NOV), 1–13. https://doi.org/10.3389/fpsyg.2016.01875

Auer, M., Malischnig, D., & Griffiths, M. (2014). Is "pop-up" messaging in online slot machine gambling effective as a responsible gambling strategy? *Journal of Gambling Issues*, (29), 1–10. https://doi.org/10.4309/jgi.2014.29.3

Blaszczynski, A., & Nower, L. (2002). A pathways model of problem and pathological gambling. *Addiction*, *97*, 487–499.

Blaszczynski, A., Russell, A., Gainsbury, S., & Hing, N. (2016). Mental health and online, land-based and mixed gamblers. *Journal of Gambling Studies*, *32*(1), 261–275. https://doi.org/10.1007/s10899-015-9528-z

Braverman, J., LaPlante, D. A., Nelson, S. E., & Shaffer, H. J. (2013). Using cross-game behavioral markers for early identification of high-risk internet gamblers. *Psychology of Addictive Behaviors*, *27*(3), 868–877. https://doi.org/10.1037/a0032818

Braverman, J., & Shaffer, H. J. (2012). How do gamblers start gambling: identifying behavioural markers for high-risk internet gambling. *European Journal of Public Health*, *22*(2), 273–278. https://doi.org/10.1093/eurpub/ckp232

Breiman, L., Friedman, J., Stone, C. J., & Olshen, R. A. (1984). *Classification and regression trees*. CRC press.

Brosowski, T., Meyer, G., & Hayer, T. (2012). Analyses of multiple types of online gambling within one provider: an extended evaluation framework of actual online gambling behaviour. *International Gambling Studies*, *12*(3), 405–419. https://doi.org/10.1080/14459795.2012.698295

Chekroud, A. M., Zotti, R. J., Shehzad, Z., Gueorguieva, R., Johnson, M. K., Trivedi, M. H., … Corlett, P. R. (2016). Cross-trial prediction of treatment outcome in depression: A machine learning approach. *The Lancet Psychiatry*, *3*(3), 243–250. https://doi.org/10.1016/S2215-0366(15)00471-X

Clark, L., & Lesch, T. (2018). *Online gambling in British Columbia: a descriptive report of gambling participation on the PlayNow.com platform*. Vancouver, Canada.

Cotte, J., & Latour, K. A. (2009). Blackjack in the Kitchen: Understanding Online versus Casino Gambling. *Journal of Consumer Research*, *35*(5), 742–758. https://doi.org/10.1086/592945

Coussement, K., & De Bock, K. W. (2013). Customer churn prediction in the online gambling industry: The beneficial effect of ensemble learning. *Journal of Business Research*, *66*(9), 1629–1636. https://doi.org/10.1016/j.jbusres.2012.12.008

Davis, J., & Goadrich, M. (2006). The relationship between Precision-Recall and ROC curves. In *Proceedings of the 23rd international conference on Machine learning (ICML '06).* (pp. 233–240). New York, USA. https://doi.org/doi.org/10.1145/1143844.1143874

Deng, X., Lesch, T., & Clark, L. (2019). Applying data science to behavioral analysis of online gambling. *Current Addiction Reports*, *6*(3), 159–164. https://doi.org/10.1007/s40429-019-00269-9

Dragicevic, S., Percy, C., Kudic, A., & Parke, J. (2015). A descriptive analysis of demographic and behavioral data from Internet gamblers and those who self-exclude from online gambling platforms. *Journal of Gambling Studies*, *31*(1), 105–132. https://doi.org/10.1007/s10899-013-9418-1

Dwyer, D. B., Falkai, P., & Koutsouleris, N. (2018). Machine learning approaches for clinical psychology and psychiatry. *Annual Review of Clinical Psychology*, *14*(1), 91–118. https://doi.org/10.1146/annurev-clinpsy-032816-045037

Edgerton, J. D., Biegun, J., & Roberts, L. W. (2016). Player behavioral tracking and personalized feedback in online gambling: implications for prevention and treatment of problem gambling. *Journal of Addiction & Prevention*, *4*(2), 1–8.

Excell, D., Bobashev, G., Gonzalez-ordonez, D., Wardle, H., Whitehead, T., Morris, R. J., & Ruddle, P. (2014). Report 3: Predicting Problem Gambling - Analysis of industry data, (November), 124.

Gainsbury, S. M. (2015). Online gambling addiction: the relationship between internet gambling and disordered gambling. *Current Addiction Reports*, *2*, 185–193. https://doi.org/10.1007/s40429-015-0057-8

Gainsbury, S. M., Abarbanel, B., & Blaszczynski, A. (2019). Factors influencing internet gamblers' use of offshore online gambling sites: policy implications. *Policy and Internet*, *11*(2), 235–253. https://doi.org/10.1002/poi3.182

Garcia, A., Dragičević, S., Percy, C., & Sarkar, S. (2019). Understanding the risk profile of gambling behaviour through machine learning predictive modelling and explanation. In

Presented at 33rd Conference on Neural Information Processing Systems (NeurIPS 2019). Vancouver, Canada. Retrieved from https://kr2ml.github.io/2019/papers/KR2ML_2019_paper_33.pdf

Gowin, J. L., Ernst, M., Ball, T., May, A. C., Sloan, M. E., Tapert, S. F., & Paulus, M. P. (2019). Using neuroimaging to predict relapse in stimulant dependence: a comparison of linear and machine learning models. *NeuroImage: Clinical*, *21*, 101676. https://doi.org/10.1016/j.nicl.2019.101676

Gray, H. M., LaPlante, D. A., & Shaffer, H. J. (2012). Behavioral characteristics of Internet gamblers who trigger corporate responsible gambling interventions. *Psychology of Addictive Behaviors*, *26*(3), 527–535. https://doi.org/10.1037/a0028545

Griffiths, M. D. (2012). Internet gambling, player protection, and social responsibility. In R. J. Williams, R. T. Wood, & J. Parke (Eds.), *The Routledge International Handbook of Internet Gambling* (pp. 227–249). London: Routledge.

Haefeli, J., Lischer, S., & Haeusler, J. (2015). Communications-based early detection of gambling-related problems in online gambling. *International Gambling Studies*, *15*(1), 23–38. https://doi.org/10.1080/14459795.2014.980297

Haeusler, J. (2016). Follow the money: using payment behaviour as predictor for future self-exclusion. *International Gambling Studies*, *16*(2), 246–262. https://doi.org/10.1080/14459795.2016.1158306

Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The elements of statistical learning: prediction, inference and data mining*. New York, USA: Springer US.

Hayer, T., & Meyer, G. (2011). Internet self-exclusion: characteristics of self-excluded gamblers and preliminary evidence for its effectiveness. *International Journal of Mental Health and*

Addiction, *9*(3), 296–307. https://doi.org/10.1007/s11469-010-9288-z

Labrie, R. A., Kaplan, S. A., Laplante, D. A., Nelson, S. E., & Shaffer, H. J. (2008). Inside the virtual casino: a prospective longitudinal study of actual Internet casino gambling. *Eur J Public Health*, *18*(4), 410–416. https://doi.org/ckn021 [pii]10.1093/eurpub/ckn021

Labrie, R. A., & Shaffer, H. J. (2011). Identifying behavioral markers of disordered internet sports gambling. *Addiction Research and Theory*, *19*(1), 56–65. https://doi.org/10.3109/16066359.2010.512106

Ladouceur, R., Jacques, C., Giroux, I., Ferland, F., & Leblond, J. (2000). Analysis of a casino's self-exclusion program. *Journal of Gambling Studies*, *16*(4), 453–460.

Langham, E., Thorne, H., Browne, M., Donaldson, P., Rose, J., & Rockloff, M. (2016). Understanding gambling related harm: A proposed definition, conceptual framework, and taxonomy of harms. *BMC Public Health*, *16*(1). https://doi.org/10.1186/s12889-016-2747-0

LaPlante, D. A., Nelson, S. E., & Gray, H. M. (2014). Breadth and depth involvement: understanding internet gambling involvement and its relationship to gambling problems. *Psychology of Addictive Behaviors*, *28*(2), 396–403. https://doi.org/10.1037/a0033810

Leino, T., Molde, H., Griffiths, M. D., Mentzoni, R. A., Sagoe, D., & Pallesen, S. (2017). Gambling behavior in alcohol-serving and non-alcohol-serving-venues: a study of electronic gaming machine players using account records. *Addiction Research and Theory*, *25*(3), 201–207. https://doi.org/10.1080/16066359.2017.1288806

Lundberg, S. M., & Lee, S.-I. (2017). A unified approach to interpreting model predictions. *Presented at 31st Conference on Neural Information Processing Systems (NeurIPS 2017)*.

Luquiens, A., Vendryes, D., Aubin, H. J., Benyamina, A., Gaiffas, S., & Bacry, E. (2018). Description and assessment of trustability of motives for self-exclusion reported by online poker gamblers in a cohort using account-based gambling data. *BMJ Open*, *8*(12), 1–8. https://doi.org/10.1136/bmjopen-2018-022541

McCormick, A. V., Cohen, I. M., & Davies, G. (2018). Differential effects of formal and informal gambling on symptoms of problem gambling during voluntary self-exclusion. *Journal of Gambling Studies*, *34*, 1013–1031. https://doi.org/10.1007/s10899-018-9743-5

Motka, F., Grüne, B., Sleczka, P., Braun, B., Örnberg, J. C., & Kraus, L. (2018). Who uses self-exclusion to regulate problem gambling? A systematic literature review. *Journal of Behavioral Addictions*, *7*(4), 903–916. https://doi.org/10.1556/2006.7.2018.96

Nelson, S. E., LaPlante, D. A., Peller, A. J., Schumann, A., LaBrie, R. A., & Shaffer, H. J. (2008). Real limits in the virtual world: self-limiting behavior of Internet gamblers. *Journal of Gambling Studies*, *24*(4), 463–477. https://doi.org/10.1007/s10899-008-9106-8

Papineau, E., Lacroix, G., Sévigny, S., Biron, J. F., Corneau-Tremblay, N., & Lemétayer, F. (2018). Assessing the differential impacts of online, mixed, and offline gambling. *International Gambling Studies*, *18*(1), 69–91. https://doi.org/10.1080/14459795.2017.1378362

Percy, C., França, M., Dragičević, S., & d'Avila Garcez, A. (2016). Predicting online gambling self-exclusion: an analysis of the performance of supervised machine learning models. *International Gambling Studies*, *16*(2), 193–210. https://doi.org/10.1080/14459795.2016.1151913

Philander, K. S. (2014). Identifying high-risk online gamblers: a comparison of data mining procedures. *International Gambling Studies*, *14*(1), 53–63. https://doi.org/10.1080/14459795.2013.841721

Philander, K. S., & MacKay, T.-L. (2014). Online gambling participation and problem gambling severity: is there a causal relationship? *International Gambling Studies*, *14*, 214–227. https://doi.org/10.1080/14459795.2014.893585

Province of BC. (2015). *A Plan for Public Health and Gambling in British Columbia*. Victoria, BC. Retrieved from https://www.gaming.gov.bc.ca/reports/docs/plan-rg-public-health-and-gambling-2015.pdf

Shaffer, H. J., Peller, A. J., LaPlante, D. a., Nelson, S. E., & LaBrie, R. a. (2010). Toward a paradigm shift in Internet gambling research: From opinion and self-report to actual behavior. *Addiction Research & Theory*, *18*(3), 270–283. https://doi.org/10.3109/16066350902777974

Strobl, C., Boulesteix, A. L., Zeileis, A., & Hothorn, T. (2007). Bias in random forest variable importance measures: Illustrations, sources and a solution. *BMC Bioinformatics*, *8*, 25. https://doi.org/10.1186/1471-2105-8-25

Sun, Y., Wong, A. K. C., & Kamel, M. S. (2009). Classification of imbalanced data: a review. *International Journal of Pattern Recognition and Artificial Intelligence*, *23*(4), 687–719. https://doi.org/10.1142/S0218001409007326

Tom, M. A., LaPlante, D. A., & Shaffer, H. J. (2014). Does Pareto Rule Internet Gambling? Problems Among The "Vital Few" & "Trivial Many." *Journal of Gambling Business and Economics*, *8*(1), 73–100. Retrieved from https://ideas.repec.org/a/buc/jgbeco/v8y2014i1p73-100.html

Wohl, M. J. A., Davis, C. G., & Hollingshead, S. J. (2017). How much have you won or lost? Personalized behavioral feedback about gambling expenditures regulates play. *Computers in Human Behavior*, *70*, 437–445. https://doi.org/10.1016/j.chb.2017.01.025

Wohl, M. J. A., Tabri, N., & Zelenski, J. M. (2019). The need for open science practices and well- conducted replications in the field of gambling studies. *International Gambling Studies*, *19*(3), 369–376. https://doi.org/10.1080/14459795.2019.1672769

Wood, R. T. A., & Wohl, M. J. A. (2015). Assessing the effectiveness of a responsible gambling behavioural feedback tool for reducing the gambling expenditure of at-risk players. *International Gambling Studies*, *15*(2), 1–16. https://doi.org/10.1080/14459795.2015.1049191

Xuan, Z., & Shaffer, H. (2009). How do gamblers end gambling: longitudinal analysis of Internet gambling behaviors prior to account closure due to gambling related problems. *Journal of Gambling Studies*, *25*(2), 239–252. https://doi.org/10.1007/s10899-009-9118-z

Author Bios

Stephanie Finkenwirth is a data scientist at the Centre for Gambling Research at UBC. She has an MSc in Econometrics from Humboldt University of Berlin, and data science experience from the private sector.

Kent MacDonald worked on this project in his role as a research assistant at the Centre for Gambling Research at UBC, and is software developer in Vancouver. He has a BA in Computer Science and Psychology from the University of Victoria.

Xiaolei Deng is a PhD student in Clinical Psychology in the Department of Psychology at the University of British Columbia. He holds a BA and MA in Psychology from UBC. His graduate research examines the use of behavioural tracking in predicting high-risk online gambling.

Tilman Lesch is a senior data scientist and manager at Deloitte Consulting GmbH in Berlin. He has a PhD in Psychiatry from the University of Cambridge, and initiated this project as part of a post-doctoral placement at the Centre for Gambling Research at UBC.

Luke Clark is a Professor in the Department of Psychology at the University of British Columbia, and Director of the Centre for Gambling Research at UBC. He is an experimental psychologist by training. His research focusses on the psychological and neural basis of decision-making in gambling and its relevance to the development of gambling harms.

Supplementary Material

Using machine learning to predict self-exclusion status in online gamblers on the PlayNow.com platform in British Columbia

Supplementary Table 1: Averages and standard deviations (in parentheses) for the whole sample, self-excluders (n=2,157) and control (n=17,526) groups. Italicized variable names reflect the 6 variables identified by the feature selection algorithm.

	Overall	Self-Excluders	Control
Total Sessions	101.7 (169.97)	90.06 (132.51)	103.13 (173.98)
Total Bets	29,193.69 (129,199.09)	30,540.66 (56,499.64)	29,027.91 (135,476.78)
Bets per Session	224.45 (253.55)	326.34 (302.3)	211.91 (243.97)

Variance in Bets per Session	126,907.9 (4,039,190.25)	177,205.18 (459,343.09)	120,717.60 (4,277,479.32)
Distinct Games per Session	2.96 (2.37)	3.35 (2.53)	2.91 (2.34)
Variance in Distinct Games per Session	9.48 (28.64)	12.22 (27.62)	9.14 (28.74)
Total Money Bet	67,103.31 (270,165.64)	85,305.28 (206,068.31)	64,863.11 (276,953.42)
Money Bet per Session	697.61 (2,315.89)	1,346.37 (3,516.15)	617.77 (2,108.16)
Variance in Money Bet per Session	12,782,508.14 (201,573,524.16)	30,631,357.37 (363,815,474.63)	10,585,773.7 (171,187,543.80)
Days Gambled	57.46 (71.82)	47.93 (55.70)	58.64 (73.47)
Sessions per Day	1.17 (0.48)	1.28 (0.57)	1.16 (0.46)

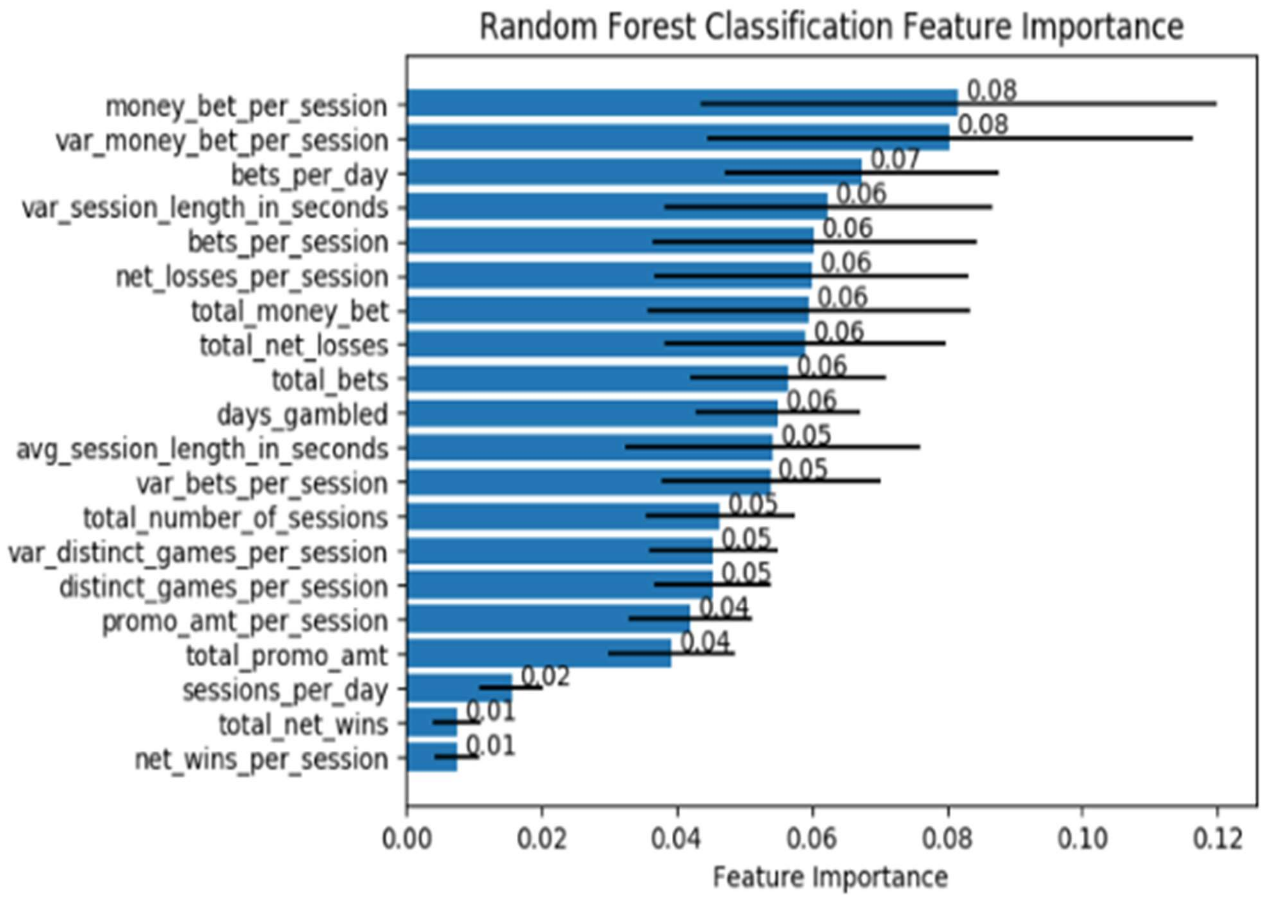
Bets per Day	345.44 (490.39)	539.16 (496.65)	321.60 (484.30)
Average Session Length	2,202.94 (1,782.58)	3010.74 (2121.08)	2,103.52 (1,710.26)
Variance in Average Session Length	7,598,110.37 (15,263,830.93)	12354395.57 (18161581.55)	7,012,733.95 (14,763,293.06)
Total Money Bet from Promotional Offers	71.61 (267.95)	78.65 (226.35)	70.74 (272.63)
Promotional Bets per Session	0.76 (2.39)	0.84 (2.63)	0.75 (2.36)
Total Net Loss	2,939.31 (9,019.51)	3926.07 (8939.18)	2,817.86 (9,022.14)
Net Loss per Session	37.72 (109.74)	72.20 (170.18)	33.48 (98.99)
Total Net Win	237.92 (6,414.07)	279.73 (4681.33)	232.78 (6,596.03)

Net Win per Session	3.39 (68.13)	5.41 (52.57)	3.14 (69.80)

Note: A high degree of skew is evidenced by the fact that in normally distributed data, most values are found within two standard deviations of the mean; for positively valued covariates, standard deviations greater than the value of the mean indicate a moderate to high degree of (positive) skew, with a minority of very high values.

Supplementary Figure 1: Correlation matrix of all 20 variables.

Supplementary Figure 2: Feature importance values from the random forest classifier from the primary model with full set of 20 input variables



Supplementary Figure 3: Histogram of predicted probabilities, in the primary random forest classifier model with all 20 variables .