

How well do machine learning algorithms predict online problem gambling?

What this research is about

Online gambling has increased worldwide, including in North America and Europe. This is because it is convenient and easy for people to use online gambling platforms, coupled with the anonymous nature of online gambling. Frequent exposure to online gambling advertisements also entices people to gamble. Concerns have been raised that online gambling poses a higher risk for problem gambling.

Online gambling platforms collect data, like deposits and amount of money won and lost. These data can be used to provide feedback to people about their gambling behaviour and identify those at risk for problem gambling. Machine learning algorithms, in particular, show promise in identifying people at risk for problem gambling.

In this study, the researchers looked at rates of problem gambling for different types of online gambling. They then explored which specific behavioural and monetary tracking features and which machine learning algorithms best predict self-reported problem gambling.

What the researchers did

The researchers analyzed self-reported and account data from 1,661 people who gambled on a North American gambling website. This website offers lottery games, casino games, bingo games, and sports betting. Participants filled out the Problem Gambling Severity Index (PGSI) between April 2023 and February 2025. Only those who had gambled at least once in the 30 days before completing the PGSI were included in the study. Participants with a PGSI score of 8 or higher were identified as having

What you need to know

In this study, the researchers looked at which behavioural and monetary tracking features in online gambling and which machine learning algorithms best predict self-reported problem gambling. The researchers analyzed data from 1,661 people who gambled on a North American gambling website. They used a mix of self-reported data on problem gambling and data from participants' online gambling accounts.

Behavioural tracking features, like number of bets made and time spent gambling, predicted self-reported problem gambling better than monetary tracking features, like amounts of money won and lost. Logistic regression and random forest were the machine learning models that best predicted problem gambling. Participants experiencing problem gambling were younger in age, more likely to be male, bet more money, bet more money on casino games and less on lotteries, deposited more money into their online accounts, had more failed deposits, and gambled for longer sessions.

The findings can inform responsible gambling strategies and interventions that gambling operators can implement to reduce harm from gambling among users.

problem gambling. The account data included information on each user's age, gender, and gambling activity.

The researchers first compared rates of self-reported problem gambling for different online gambling

types. They then looked at whether the likelihood of experiencing problem gambling was best predicted by behavioural tracking features alone (e.g., time spent gambling), or behavioural and monetary tracking features (e.g., money spent on gambling) together. Finally, they tested several machine learning models to see which best predicted self-reported problem gambling.

What the researchers found

On average, participants were 55 years old. There were 693 males, 690 females, and 227 who identified as “other” (including people who did not want to disclose their gender). About 13% of participants were classified as having problem gambling based on the PGSI. Lotteries were the most common online gambling type, followed by casino games, then sports betting, and finally bingo games.

Engaging in multiple types of online gambling was linked to a greater likelihood of self-reported problem gambling. However, there were very few participants (0.5%) who engaged in all four types of online gambling. The highest rate of problem gambling was among people who only engaged in online casino games.

Participants who only engaged in lotteries were at lower risk for problem gambling. Therefore, they were excluded from the analysis that looked at predictors of problem gambling. A logistic regression model that included only behavioural tracking features, like number of bets made and time spent gambling, predicted self-reported problem gambling better than a model that also included monetary tracking features, like amounts of money won and lost. In other words, monetary tracking features did not add further value in predicting problem gambling over behavioural tracking features alone.

The researchers then tested how well five machine learning algorithms predicted problem gambling: logistic regression, decision tree, random forest, support vector machine, and gradient boosting machine. Logistic regression performed the best overall, correctly identifying 52% of participants with self-reported problem gambling. Random forest also performed better than other algorithms.

Compared to people not experiencing problem gambling, people experiencing problem gambling tended to be younger in age, more likely to be male, bet more money, bet more money on casino games and less on lotteries, deposited more money into their online accounts, had more failed deposits, and gambled for longer sessions.

How you can use this research

The findings of this study provide insight into how user data and machine learning algorithms can be leveraged to predict online problem gambling. They can inform responsible gambling strategies and interventions that gambling operators can implement to reduce harm among users.

About the researchers

Michael Auer is affiliated with Neccton GmbH in Müllendorf, Austria. **Mark D. Griffiths** is affiliated with the Psychology Department at Nottingham Trent University in Nottingham, United Kingdom. For more information about this study, please contact Mark D. Griffiths at mark.griffiths@ntu.ac.uk.

Citation

Auer, M., & Griffiths, M. D. (2026). Using machine-learning algorithms to predict self-reported problem gambling among a sample of online gamblers. *International Journal of Mental Health and Addiction*. Advance online publication. <https://doi.org/10.1007/s11469-025-01602-2>

Study Funding

This study received no direct funding.

About Greo Evidence Insights

Greo Evidence Insights is an independent, non-profit organization specializing in research, knowledge mobilization, and evaluation in the health and well-being sectors. Greo transforms research and insights into actionable strategies to prevent and reduce harm related to gambling, gaming, technology use, and substance use. To learn more about funding for Research Snapshots visit [our website](#).

