**What this research is about**

Online gambling activities have become more popular in recent years. A concern about online gambling is that most of its revenue comes from a minority of highly engaged users. Higher participation in online gambling can potentially increase gambling-related harms. Gambling-related harms can range from individual impacts to broader familial and societal impacts. Such harms may include relationship stress and increased risks for suicide attempts.

Studies have shown that online gambling platforms have more users experiencing harms compared to land-based gambling venues. Yet, online platforms are equipped with the ability to provide personalized assessments and referrals to treatment services.

Online operators and regulators can examine people’s gambling behaviour in depth. This could allow them to develop systems that are able to identify people at risk of gambling harms. Such systems could allow new forms of personalized support to emerge on online gambling platforms. This idea has been well received by people who gamble, researchers, and professionals. To this end, more gambling researchers are applying machine learning. Machine learning is a powerful tool for analyzing large datasets in attempts to identify people who are at risk of harms. It is used to make predictions based on indicators of their gambling behaviour.

The current study uses the Problem Gambling Severity Index (PGSI) as an indicator of past-year gambling harms. The researchers used machine learning to classify people who are at risk of harms using data generated from their use of an online gambling website.

**What you need to know**

Online gambling is becoming more and more popular. Research has shown that data collected by online gambling platforms can be used to give personalized assessment or referrals to treatment services for users who experience harm. The researchers used the Problem Gambling Severity Index (PGSI) and machine learning to classify people who are at higher risk of online problem gambling. Using data collected from people registered on lotoquebec.com, the researchers found that machine learning algorithms could be used with online gambling data to predict whether a person belongs to a PGSI higher-risk category.

**What the researchers did**

People registered on lotoquebec.com were invited to complete a survey about their gambling habits. lotoquebec.com is an official site offering online gambling games in Quebec, Canada. The final sample included 9,145 participants. Among them, 88 participants were excluded because they did not place a bet in the 12 months prior to the survey.

The survey was offered both in French and in English. It included the PGSI, which has 9 items rated from 0 (Never) to 3 (Almost always). The validated risk categories for the PGSI are as follows:

- Non-problem or low-risk gambling (scores 0–4);
- Moderate-risk gambling (scores 5–7);
- High-risk of problem gambling (scores 8–27).

The researchers defined two dependent variables (i.e., outcomes) of interest: PGSI 5+ and PGSI 8+. The
PGSI 5+ included people who scored as moderate- to high-risk of problem gambling. The PGSI 8+ included people with high-risk of problem gambling.

The researchers generated predictors using the data collected on lotoquebec.com over the 52 weeks before participants completed the PGSI. A total of 10 predictor variables were included. The predictors included demographic information, indicators of online gambling behaviour, and repeat engagement.

The researchers used six machine learning algorithms and determined which one would generate the best model in predicting PGSI 5+ and PGSI 8+. The algorithms included logistic regressions, decision trees, K-nearest neighbours, support vector machines, neural networks, and random forests.

What the researchers found

The results showed that participants with higher risk of gambling problems (PGSI 5+ or PGSI 8+) were significantly younger. Moreover, females were over-represented in these higher-risk groups.

All six machine learning algorithms generated good classification performance. The random forest approach was selected as the best among them. The analysis showed that the PGSI 5+ model based on the random forest approach correctly classified 81.75% of the participants with PGSI 5+ and 74.36% of the participants with PGSI < 5. The PGSI 8+ model correctly classified 81.94% of the participants with PGSI 8+ and 72.20% of the participants with PGSI < 8.

However, both models demonstrated low precision. Precision referred to the proportion of positive predictions correctly made by a model regarding higher-risk gambling. The PGSI 5+ model indicated a 46.29% precision. The PGSI 8+ model demonstrated a 29.48% precision. This suggests that most participants who were predicted to be at higher risk actually belonged to the lower-risk group.

A few predictors were found to be important in predicting whether participants had higher risk (PGSI 5+ or PGSI 8+). These included maximum bets, repeat engagement, and variation in weekly bet amounts.

How you can use this research

This research can be used by online gambling sites that are interested in implementing machine learning. It shows that machine learning algorithms can be used with online gambling data to predict PGSI at-risk categories. The information can then be used to provide personalized support and referrals.

About the researchers

W. Spencer Murch, Sylvia Kairouz, Sophie Dauphinais, Elyse Picard, Jean-Michel Costes, and Martin French are affiliated with the Department of Sociology and Anthropology at Concordia University in Montreal, Quebec, Canada. For more information about this study, please contact Sylvia Kairouz at sylvia.kairouz@concordia.ca.

Citation


Study funding

This study was funded by a Concordia University Horizon Postdoctoral Fellowship, the Mise sur toi Foundation, and the Fonds de recherche du Québec – Société et Culture.

About Greo

Greo has partnered with the Knowledge Mobilization Unit at York University to produce Research Snapshots. Greo is an independent knowledge translation and exchange organization with almost two decades of international experience in generating, synthesizing, and mobilizing research into action across the health and wellbeing sectors. Greo helps organizations improve their strategies, policies, and practices by harnessing the power of evidence and stakeholder insight.

Learn more about Greo by visiting greo.ca or emailing info@greo.ca.