Facebook posts may reveal individuals at risk for excessive drinking

May 20 2022

This visual presentation shows the words and phrases used on Facebook posts from individuals in the study considered either high- or low-risk for excessive drinking. Credit: Rupa Jose and Andrew Schwartz

Alcoholism can be a difficult condition to diagnose, especially in cases where individuals' drinking habits are not noticed and physical symptoms have not yet manifested. In a new study, published in *Alcoholism: Clinical & Experimental Research*, co-author H. Andrew Schwartz, Ph.D., of the Department of Computer Science at Stony Brook University, and colleagues determined that the language people used in Facebook posts can identify those at risk for hazardous drinking habits and alcohol use disorders.

Collaborating with Schwartz working on The Data Science for Unhealthy Drinking Project is Stony Brook University doctoral candidate Matthew Matero, and Rupa Jose, Ph.D., lead author and Postdoctoral Researcher at the University of Pennsylvania.

Key to the research was the use of Facebook content analyzed with "contextual embeddings," a new artificial intelligence application that interprets language in context. The contextual embedding model, say Schwartz, Jose and colleagues, had a 75 percent chance of correctly identifying individuals as high- or low-risk drinkers from their Facebook posts. This rate at identifying at risk people for excessive drinking is higher than other more traditional models that identify high-risk drinkers and those vulnerable to alcoholism.

"What people write on social media and online offers a window into psychological mechanisms that are difficult to capture in research or medicine otherwise," says Schwartz, commenting on the unique aspect of the study.

"Our findings imply that drinking is not only an individually motivated behavior but a contextual one; with social activities and group membership helping set the tone when it comes to encouraging or discouraging drinking," summarizes Jose.

Investigators used data from more than 3,600 adults recruited online—average age 43, mostly White—who consented to sharing their Facebook data. The participants filled out surveys on demographics, their drinking behaviors, and their own perceived stress, a risk factor for problematic alcohol use. Researchers then used a diagnostic scale to organize participants, based on their self-reported alcohol use, into high-risk drinkers (27 percent) and low-risk drinkers (73 percent).

The Facebook language and topics associated with high-risk drinking included more frequent references to going out and/or drinking (e.g., "party," "beer"), more swearing, more informality and slang ("lmao"), and more references to negative emotions ("miss," "hate," "lost," and "hell"). These may reflect factors associated with high-risk drinking, including neighborhood access to bars, and personality traits such as impulsivity.

Low-risk drinking status was associated with religious language ("prayer," "Jesus"), references to relationships ("family," "those who"), and future-oriented verbs ("will," "hope"). These may reflect meaningful support networks that encourage drinking moderation and the presence of future goals, both of which are protective against dangerous drinking.

Overall, the authors conclude that "social media data serves as a readily available, rich, and under-tapped resource to understand important public health problems, including excessive alcohol use...(The) study findings support the use of Facebook language to help identify probable alcohol vulnerable populations in need of follow-up assessments or interventions, and note multiple language markers that describe individuals in high/low alcohol risk groups."

Provided by Stony Brook University