From Words to Actions: A Comprehensive Approach to Identifying Incel Behavior on Reddit

Demirbas, Ahmet (School: Williamsville East High School)

The incel (involuntary celibate) community is a radicalized online subculture. In this study, I focus on the popular social media platform Reddit and analyze the activity patterns and comments of 14,000 users to characterize the mechanics of online radicalization and offer tools for identifying at-risk individuals before they become deeply entrenched in incel ideologies. My work utilizes systematic data collection to construct a 7000 user incel-positive group and an equal-sized incel-negative group who share similar activity levels in the neutral subreddits frequented by the positive users. I devise a hybrid method that combines both language analysis on users' comments and their behavior patterns across different subreddits. I find that utilizing language analysis alone yields a classification accuracy of 76%, while using only the subreddit activity patterns of users achieves an accuracy of 79%. Remarkably, my hybrid method blends both approaches and reduces false positives and negatives by cross-referencing data to achieve 99.8% accuracy, improving 20% over the state-of-the-art classification solutions. This work is important because it can spot individuals at risk of joining harmful groups early on, allowing for timely support. While focused on Reddit and incels specifically, this hybrid method can be adapted across other online communities and platforms, and the insights from this work can guide platform moderators and policymakers in creating safer online spaces.

Awards Won: