Rush Weigelt

Senior Design Project Proposal

Twitter/4Chan Study and Automation of ML Processes

9/9/19

**Twitter/4Chan Study and Automation of ML Processes**

**Abstract:**

While not universal, many mass shooters in the United States leave a vast online footprint in the form of writing—whether that is on Twitter, Facebook, or a forum like 4Chan. We propose using these writings to train a supervised learning model to help categorize users as “potentially dangerous” or not, thus automating the “see-something, say-something” defense mechanism currently in place.

The current model of “see something, say something,” has been responsible for a recent rash of threat-related arrests across the US; however, this passive defense mechanism is prone to various holes as it relates to social media. Whether it is because the User’s online reach is relatively weak, their close online associates hesitate to notify authorities, or they don’t make any direct threats towards any other account, it is easy to miss what, in retrospect, becomes an obvious warning sign.

**Proposal:**

First, test out various models to achieve a desirable accuracy score for supervised learning into two categories: ‘potentially dangerous’ and ‘benign’. Simultaneously, start learning how to properly webscrap Twitter posts and forum (4Chan) posts in mass quantities. Next, automate both processes to run on a clickable button. Lastly, export all categorized posts/Users into a single CSV file, with clickable links for a human user to sift through.