Developing a Twitter 'Bot' Identification Application for Public Use

Rilatt, Adam (School: Central York High School)
Feistritzer, Daniella (School: Central York High School)

The goal for this project was to create a user-friendly program capable of accurately detecting ‘bots’ (automated accounts) on Twitter, inspired by the influx of bot activity and public concern following the 2016 United States presidential election. It builds on results from the previous year’s project. Researchers trained a machine learning model on 2,000 Twitter accounts confirmed as bot or human, using 10 input features selected from data provided via the Twitter API. To increase public accessibility, researchers developed a Chrome extension which scrapes usernames from a Twitter feed, sends them to a locally-hosted Web application, collects input features via the Twitter API, runs the input features through the machine learning model to generate the probabilities that the accounts are bots, and returns the probabilities to the Chrome extension to be displayed next to the usernames. The final result was a machine learning model that demonstrated a 93.75% accuracy on a 400-account test set. The algorithm was more accurate labeling bots than humans, which researchers attributed to the complexity of human interactions online. The client-server model exceeded the researchers’ expectations of 85% accuracy. This project will increase user awareness of bot presence and their influence on the user experience, therefore promoting a safer online experience.

Awards Won: