UC-62 Machine Learning: Twitter Bots in Disguise

Abstract

Today's popular social media platforms such as Instagram, Twitter, and Facebook are filled with computer operated accounts (bots). Users of these social media platforms may not even know that their accounts could be followed or they could be following these fake accounts. While some are not harmful and mostly spam accounts, there still exists bot accounts that raise privacy concerns to users everywhere.

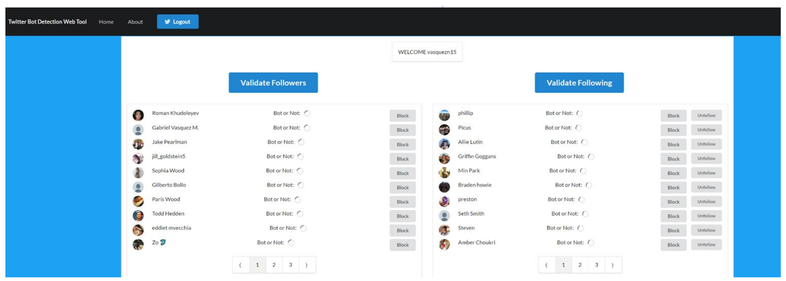
This application aims to inform Twitter users on exactly how many accounts that follow them or that they follow are bot accounts with 80% accuracy using machine learning. The problem is finding the best machine learning model for this solution. The team tested several different models on data sets with accounts confirmed to be bots and datasets with accounts confirmed to be humans,

The results show that K-Nearest Neighbors model might be the best solution for this application, boasting 90-92% accuracy through SKlearn metrics.

Introduction

There have been an increasing number of Twitter bots within the app and that number greatly increased under the COVID-19 pandemic. Not all bots are harmful but some of them are and can fill sites with spam, spread misinformation, and attempt to scam users, among various other nefarious actions. Our goal is to give users the opportunity to step away from bots, evaluate their social media presence, and better understand potential "tells" of bot accounts.

The Twitter Bot Detector web tool allows users to assess their account's followership and friends to determine which accounts are bots. The tool is intended for Twitter users who may be concerned about their privacy, security, or peace of mind regarding the presence of bots on the twitter website, and allows the user to unfollow or block accounts identified as bots.



Above: The web app's interface. Allows user to block or unfollow analyzed accounts.

Research Question(s)

1. How do we distinguish bots from real people?
2. What machine learning algorithms are most viable for distinguishing bots from people?
3. How can we be more aware of bots on social media platforms?

Materials and Methods

Development Tools:

Using some already-available datasets online, we developed models to identify bots based on certain criteria.

These signifiers include criteria such as:

- Follower and friend count
- Verified status
- Location set in bio
- If the user has a background image
- If the user uses the default profile icon image

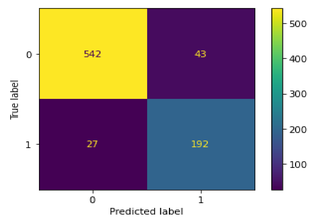
We trained several different models (Naive bayes, K-nearest neighbors, and support vector machine) with sklearn, on the same datasets. Afterward, we compared them on accuracy, precision, recall, and F1 score to determine the best algorithm for the purposes of this analysis. To save resources and avoid recomputation, the trained models were serialized for further reference.

From the web app, users grant the tool access to their followers and friends lists (through Twitter's API); these user ids are passed to the server, which analyzes their profiles and tweet trends and returns a classification for each account.

Results

Machine Learning

Among the algorithms we tested for this project, K-nearest neighbors performed the best for bot classification, producing 91% accuracy and 88% precision on the validation dataset. The level of precision indicates a greater proportion of false positives (i.e. marked as bots) than desired; in exploring our datasets, we found that many accounts labeled "human" were verified, likely leading to the algorithm determining that *non-verified* accounts are also non-human. This was caught and corrected during future passes, and different datasets with more representative samples of verified, non-verified, human, and bot accounts were used instead. To mitigate this issue in the future, larger datasets can be used, or an alternate algorithm such as support vector machine can be set to weight the criteria less; the probability threshold can also be adjusted to improve metrics.



Left: Confusion matrix for KNN

Twitter Access

Twitter's API requires specific authorization requirements to access user sensitive data. The process for encoding a request can be complex, so an additional framework to authenticate our requests was required. After authenticating requests, a problem may occur where a request can reach a limit, which will terminate the application. To bypass this, we utilized React's framework to make the request one time, and store it within a state variable of the application. This allows the app to make fewer requests to the API, avoiding the rate limit.

Conclusions

Our results demonstrate that machine learning models can identify social media bots with reasonable accuracy. However, the models in this project had difficulty with misidentifying human accounts as bots. Similar issues were reported for tools like BotOrNot (see the second paper below); one of the author's accounts was also misidentified as a bot by Botometer, a more widely used app. This suggests issues in generalizing human accounts by profile information alone -- for example, in Internet parlance, the term "lurker" refers to a user who rarely posts or tweets; in terms of algorithms, this behavior would likely point to a bot.

It is important to consider the objective of bot classification as well. Judgments about the value of *finding as many bots as possible* (at the cost of misidentifying human accounts) versus minimizing misclassification of human users is likely to be left to stakeholders (Twitter itself and its users).

Further study may serve to clarify this issue through identification of how potentially "dangerous" or threatening an account may be. Criteria such as the number of impressions the account makes on others or on popular discourse, how frequently the account posts or is active, and whether the account identifies itself as a bot may be useful in this area.

Acknowledgments

Dr. Reza Parizi rparizi1@kennesaw.edu

Contact Information

James Andersen jande276@students.kennesaw.edu
Daniel Rimmel drimmel@students.kennesaw.edu
Matthew Scheer mscheer1@students.kennesaw.edu
Josh Tiangco jtiangco@students.kennesaw.edu
Justin Van jvan@students.kennesaw.edu
Nicolas Vasquez nvasque2@students.kennesaw.edu
Cody Walicek cwalicek@students.kennesaw.edu

References

Supervised Machine Learning Bot Detection Techniques to Identify Social Twitter Bots - academic paper
- https://scholar.smu.edu/cgi/viewcontent.cgi?article=1019&context=datasciencereview

An in-depth characterisation of Bots and Humans on Twitter - academic paper
- https://arxiv.org/pdf/1704.01508.pdf

Twitter Development Kit
- https://developer.twitter.com/en/docs/api-reference-index

Botometer - Similar bot detection tool
- https://botometer.osome.iu.edu/

KENNESAW STATE UNIVERSITY

Author(s): James Andersen, Daniel Rimmel, Matthew Scheer, Josh Tiangco, Justin Van, Nico Vasquez, Cody Walicek
Advisor: Dr. Reza Parizi