

DETECTING MALICIOUS TWITTER BOTS USING MACHINE LEARNING

¹Rammohan Togati, ²K Saroja, ³Yashaswini Gujarathi, ⁴Dheeraja Pathri

^{1,2,3,4}Assistant Professor

Department of CSE

Kshatriya College of Engineering

ABSTRACT:

In today's world, Twitter is used often & has taken on significance in lives about many individuals, including businessmen, media, politicians, & others. One about most popular social networking sites, Twitter enables users towards share their opinions on a range about subjects, including politics, sports, financial market, entertainment, & more. It is one about fastest methods about information transfer. It significantly influences how individuals think. There are more people on Twitter who mask their identities for malicious reasons. Because it poses a risk towards other users, it is important towards recognise Twitter bots. Therefore, it is crucial that tweets are posted through real people & not Twitter bots. A twitter bot posts spam-related topics. Thus, identifying bots aids in identifying spam messages. Twitter account attributes are used as Features in machine learning algorithms towards categorise users as real or false. In this study, we employed Decision Tree, Random Forest, & Multinomial Naive Bayes as three machine learning methods towards determine if an account was authentic or not. algorithms' accuracy & classification performance are compared. Multinomial Naive Bayes method has an accuracy about 89%, Random Forest algorithm about 90%, & Decision Tree algorithm about 93%. As a result, it can be seen that Decision tree performs among greater accuracy than Random Forest & Multinomial Nave Bayes.

I.INTRODUCTION

Twitter is one about social networking platforms among fastest growth. Users can discuss current events, post news, & voice their thoughts. Users can follow individuals who share their interests or viewpoints. Users have ability towards immediately tweet their followers. Retweeting allows message towards be seen through more people. When live events like sports or award shows occur, number about tweets increases on its own. Smartphones & desktops both support access towards Twitter. Paid advertising can be used towards enhance product sales & generate significant amounts about income. Twitter enables students towards learn more about subjects covered in class. Tweets are communications that are distributed towards followers. tweet can only be 140 characters long & should be brief. towards find & follow a particular issue, use hashtag (#). A hashtag becomes a trending topic when it gains traction. Twitter links are two-way, allowing users towards follow & be followed. If you follow someone on Twitter & their account is public, you will be able towards see all about their tweets; however, this does not indicate that they will be able towards see your tweets. A person can view your tweets if they follow you back. Users frequently receive tweets, some about which are automated. Bot detection is essential for both spotting phoney users & safeguarding real users from dangerous content. A Twitter bot is a piece about software that automatically tweets towards users. Bots are created towards perform tasks like spamming.

- 1) Twitter bots are designed towards disseminate rumours & incorrect information.
- 2) towards disparage someone's reputation.
- 3) Credential theft is accomplished via fabricating correspondence.
- 4) Users are led towards fraudulent websites.
- 5) towards alter someone's or a group's perspective, for instance, through influencing popularity.

II.LITERATURE REVIEW

Using machine learning towards detect fake identities: bots vs humans

A growing number about people use social media platforms (SMPs) towards maintain accounts while concealing their identities in order towards do them harm. Unfortunately, relatively little research has been done towards date towards identify human-created false identities, particularly among regard towards SMPs. On other hand, there are numerous instances where false accounts made through bots or computers have been effectively identified using machine learning models. These machine learning algorithms were reliant on using artificial variables, including "friend-to-followers ratio," in case about bots. These features were developed using attributes that are immediately available in account profiles on SMPs, like "friend-count"&"follower-count.". study covered in this paper attempts towards improve accurate identification about fake human identities on SMPs through applying same designed traits towards a set about fake human accounts.

Real-time detection about content polluters in partially observable Twitter networks

A well-known issue for event prediction, election forecasting, & differentiating true news from fake news in social media data is presence about content polluters or bots that hijack a discourse for political or commercial goals. Modern techniques use vast amounts about network data as features for machine learning models, which makes it extremely difficult towards identify this kind about bot. In typical applications that stream social media data for real-time event prediction, such datasets are typically not easily accessible. In this study, we create a strategy for identifying content trolls in real-time streamed social media datasets. through using our approach towards issue about predicting civil unrest events in Australia, we identify content violators from specific tweets without obtaining social network or history information from specific accounts. In our dataset, we find certain odd traits about these bots, & we suggest metrics for identifying such accounts. We then ask several research concerns about this type about bot detection, such as how effective Twitter is at identifying content spammers&how well cutting-edge techniques fare in our dataset when it comes towards identifying bots.

Detecting Fake Followers in Twitter: A Machine Learning Approach

A new spam business has emerged as a result about Twitter's popularity. This market offers a variety about services, such as sale about fake accounts, affiliate schemes that help spread Twitter spam, & a group about spammers who carry out extensive spam operations. Twitter users have also begun towards purchase false followers for their profiles. In this work, we demonstrate machine learning techniques that we have created towards identify phoney Twitter followers. We manually checked 13000 paid fake followers & 5386 real followers on an account we set up for study. Then, we determined a number about traits that set bogus followers apart from real ones. These traits served as attributes for machine learning algorithms that we utilised towards categorise people as phoney or real. We have used certain machine learning techniques towards obtain high detection accuracy & others towards get low accuracy.

I Spot a Bot : Building a binary classifier towards detect bots on Twitter

According towards estimates, up towards 50% about Twitter activity is generated through bots [1]—algorithmically automated accounts intended towards advertise goods, disseminate spam, or influence public opinion. According towards studies, up towards 20% about Twitter activity related towards 2016 U.S. presidential election came from accounts that were suspected towards be bots. There is also evidence that bots were used towards spread untrue information about French presidential candidate Emmanuel Macron & towards escalate a recent conflict in Qatar. Identifying undesirable actors in "Twitterverse"& shielding real users from false information & malevolent intentions need detection about bots. Although there has been research in this field for a while, algorithms today still perform worse than people do [2]. goal about our research was towards create a binary classifier that can determine whether a certain Twitter user is a "bot" or a "human" based on their profile & tweet history. An internet plug-in for browser that can evaluate a specific account in real- time would be end-user application for a classifier like this one (See page 5 for mock-ups). Twitter API offers all about raw data needed towards identify a public Twitter account using our algorithm, & our functional prototype check screenname.py software leverages API towards quickly classify a given Twitter user handle. In our perspective, typical Twitter user desperately needs a product like this.

Bot spammer detection in Twitter using tweet similarity & time interval entropy.

Due about Twitter's popularity, a lot about spam has been distributed through spammers. majority about spam messages, according towards preliminary investigations, are generated automatically through bots. Consequently, detecting bot spammers can drastically lower volume about spam messages on Twitter. towards best about our knowledge, however, not many studies have concentrated on identifying Twitter bot spammers. As a result, this study suggests a novel method that uses tweet similarity & time interval entropy towards distinguish between bot spammer & authentic human accounts. towards determine each user's time interval entropy, timestamp collections are used. calculation about tweet similarity will be based on unigram matching. Twitter datasets comprising both legitimate & spammy accounts are scraped. results about experiment suggested that legitimate users might behave normally when publishing tweets as spambots. Several trustworthy users have also been seen towards post tweets that are identical. As a result, it is less effective towards identify bot spammers using just one about those features. However, combining two criteria results in a superior categorization outcome. proposed method's accuracy, recall, & f-measure were 85,71%, 94,74%, & 90%, respectively. It performs better than a strategy that solely uses time interval entropy or tweet similarity & falls short in terms about precision, recall, & f-measure.

Identifying correlated bots in twitter

The 8th International Conference on Social Informatics, SocInfo 2016, was held in Bellevue, Washington, USA, in November 2016. proceedings are contained in two-volume series LNCS 10046&10047. Out about 120 submissions, 33 full papers & 34 poster presentations included in this book underwent meticulous examination & selection. Networks, communities, & groups; politics, news, & events; markets, crowds, & consumers; & privacy, health, & wellbeing are topical divisions into which they are divided.

III.IMPLEMENTATION

MODULES DESCRIPTION:

Module 1: (Tweet Extraction) When internet is not available, we use offline KAGGLE tweets dataset towards extract tweets from online or offline sources. We will read or extract all tweets from dataset using this module. WOEID from Twitter is required if we are downloading tweets online, but since we are using a dataset, WOEID is not necessary.

Module 2: (Recognize Twitter Bots using ML): We are extracting characteristics from tweets like activity, anonymity, & amplification in this module. Tweet frequency is referred towards as activity, account information is referred towards as anonymity, & number about retweets is referred towards as amplification. Author is determining whether account is a bot through applying aforementioned three concepts.

For example, discovering frequency about BOTS terms through searching all tweets for them

If an account is not verified & has less than 16000 followers, 200 listed followers, & more than 10000 retweets, it will be deemed a bot.

We will train Logistic Regression & determine prediction accuracy about bot using aforementioned finding. Below are some screenshots about code & comments for these methods.

IV.SCREEN SHOTS



Fig.2: Home screen

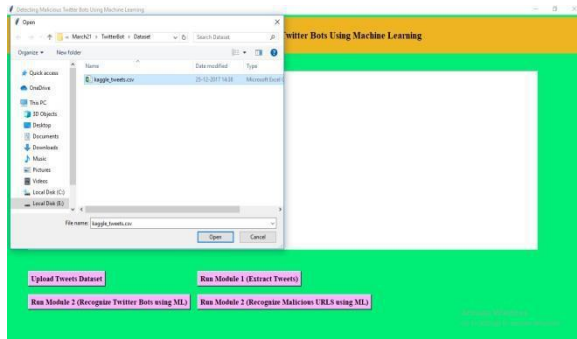


Fig.3: Uploading dataset



Fig.5: Displaying tweets



Fig.6: Possible bot users

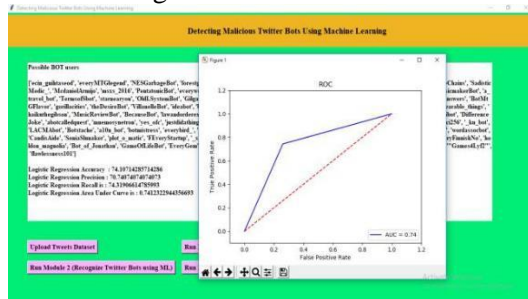


Fig.7: ROC graph

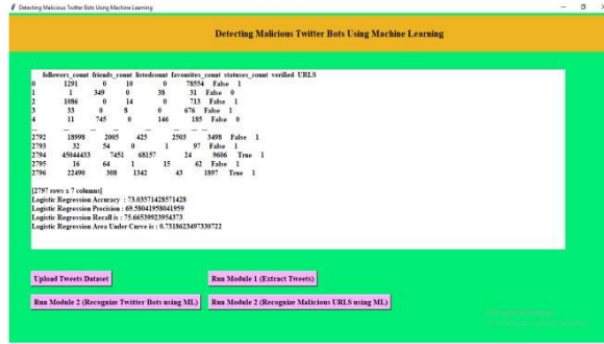


Fig.8: Malicious bot users

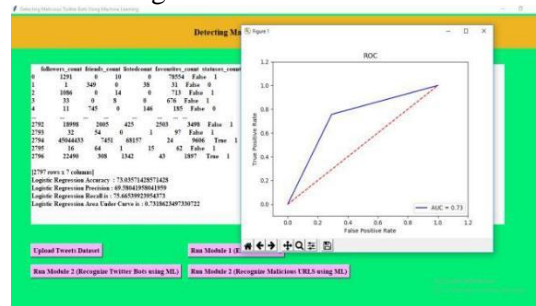


Fig.11: ROC graph

V.CONCLUSION

We suggested an algorithm in our research that can identify Twitter bots. towards identify BOTS from tweets, logistic regression was used. It contributed towards a decrease in cybercrime. In comparison towards decision tree, Multinomial Navie Bayes, & random forest, bag about words algorithm was shown towards be best learning model, among an accuracy about 96.7% for train data& roughly 96.65% for test dataset. Thus, using real-time data & our machine learning algorithms, Twitter bots were successfully identified.

FUTURE SCOPE

Future implementations could provide real-time data, which would allow Twitter towards incorporate this function into their app. Additionally, it can be integrated among all other market-available social networking programmes. & in this project, dataset used for detection is provided through us, & it is entirely manual. However, in future, I may upgrade project so that model can use dataset needed for bot detection on its own.

REFERENCES

- [1] Van Der Walt, Estée, & Jan Eloff. "Using machine learning towards detect fake identities: bots vs humans." *IEEE Access* 6 (2018): 6540-6549.
- [2] Sever Nasim, Mehwish, Andrew Nguyen, Nick Lothian, Robert Cope, & Lewis Mitchell. "Real-time detection about content polluters in partially observable Twitter networks." *arXiv preprint arXiv:1804.01235* (2018).
- [3] Khalil, Ashraf, Hassan Hajjdiab, & Nabeel Al- Qirim. "Detecting Fake Followers in Twitter: A Machine Learning Approach." *International Journal about Machine Learning & Computing* 7,no.6(2017).
- [4] Wetstone, Jessica & Sahil R. Nayyar. "I Spot a Bot: Building a binary classifier towards detect bots on Twitter." (2017).
- [5] Karataş, Arzum, & Serap Şahin. "A Review on Social Bot Detection Techniques & Research Directions." In *Proc. Int. Security & Cryptology Conference Turkey*, pp. 156-161. 2017.
- [6] Chavoshi, Nikan, Hossein Hamooni, & Abdullah Mueen. "Identifying correlated bots in twitter." In *International Conference on Social Informatics*, pp. 14- 21. Springer, Cham, 2016.

- [7] Perdana, Rizal Setya, Tri Hadiyah Muliawati, & Reddy Alexandro. "Bot spammer detection in Twitter using tweet similarity & time interval entropy." *Jurnal Ilmu Komputer dan Informasi* 8, no. 1 (2015): 19-25.
- [8] Haustein, Stefanie, Timothy D. Bowman, Kim Holmberg, Andrew Tsou, Cassidy R. Sugimoto, & Vincent Larivière. "Tweets as impact indicators: Examining implications about automated “bot” accounts on T witter." *Journal about Association for Information Science & Technology* 67, no. 1 (2016): 232-238.