Detecting Malicious Twitter Bots Using Machine Learning

K Arjun[1], GS Udaya Kiran Babu[2], M Anand[3], Syeda Fatima Sayema[4]

[1,2]Associate Professor,[3]Assistant Professor,[4]Student
Department Of CSE
Bheema Institute of Technology and Science, Adoni

ABSTRACT: Twitter is widely used and has become important in the lives of many people in the modern world, including politicians, businesspeople, and the media. Twitter is one of the most widely used social networking sites. It allows users to express their ideas on a variety of topics, such as politics, sports, the financial market, entertainment, and more. It is among the quickest ways to transport information. It has a big impact on people's thinking. On Twitter, the number of individuals hiding their identities for nefarious purposes is increasing. It's critical to identify Twitter bots since they present a danger to other users. As a result, it's critical that actual individuals, not Twitter bots, publish tweets. Spam-related topics are posted by a Twitter bot. Consequently, recognising bots helps distinguish spam communications. Machine learning algorithms analyse features from Twitter accounts to classify individuals as authentic or fake. To ascertain if an account was real or not, we used three machine learning techniques in this study: decision trees, random forests, and multinomial naive bayes. The accuracy and classification performance of the algorithms are compared. About 89% of decisions are made correctly by the Multinomial Naive Bayes technique, 90% by the Random Forest algorithm, and 93% by the Decision Tree algorithm. Consequently, it is evident that decision trees outperform Random Forest and Multinomial Nave Bayes in terms of accuracy.

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

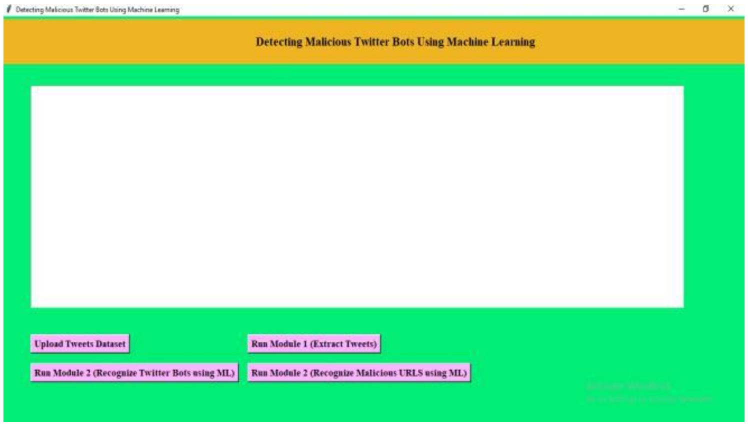

Detecting Malicious Twitter Bots Using Machine Learning
Detecting Malicious Twitter Bots Using Machine Learning
Upload Tweets Dataset
Run Module 1 (Extract Tweets)
Run Module 2 (Recognize Twitter Bots using ML)
Run Module 2 (Recognize Malicious URLS using ML)

Fig.2: Home screen

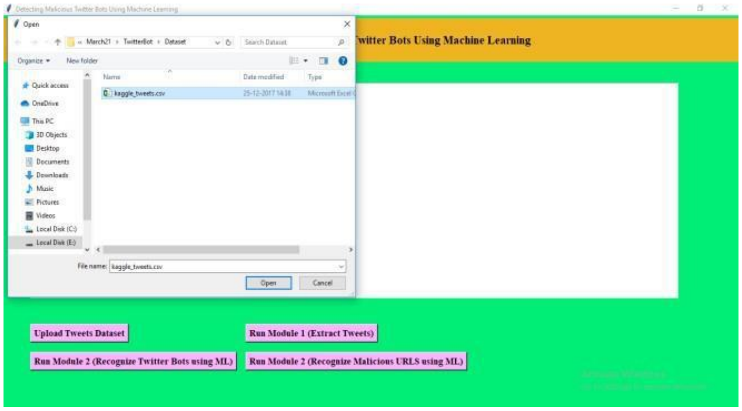
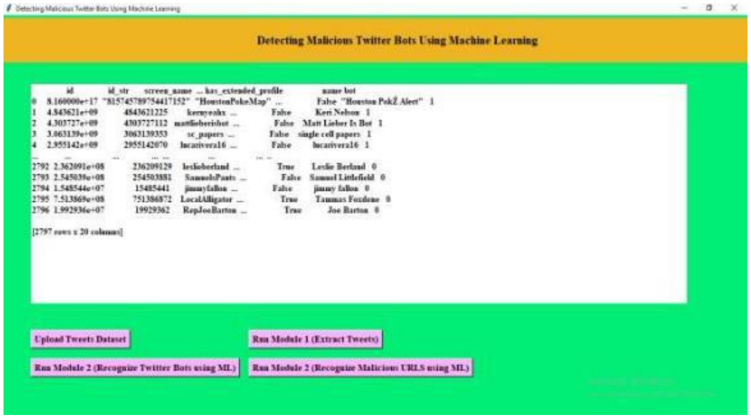

Detecting Malicious Twitter Bots Using Machine Learning
Open
March21 > TwitterBot > Dataset
Search Dataset
Organize
New folder
Quick access
OneDrive
This PC
3D Objects
Desktop
Documents
Downloads
Music
Pictures
Videos
Local Disk (C:)
Local Disk (E:)
Name
Date modified
Type
kaggle_tweets.csv
25-12-2017 14:38
Microsoft Excel
File name: kaggle_tweets.csv
Open
Cancel
witter Bots Using Machine Learning
Upload Tweets Dataset
Run Module 1 (Extract Tweets)
Run Module 2 (Recognize Twitter Bots using ML)
Run Module 2 (Recognize Malicious URLS using ML)

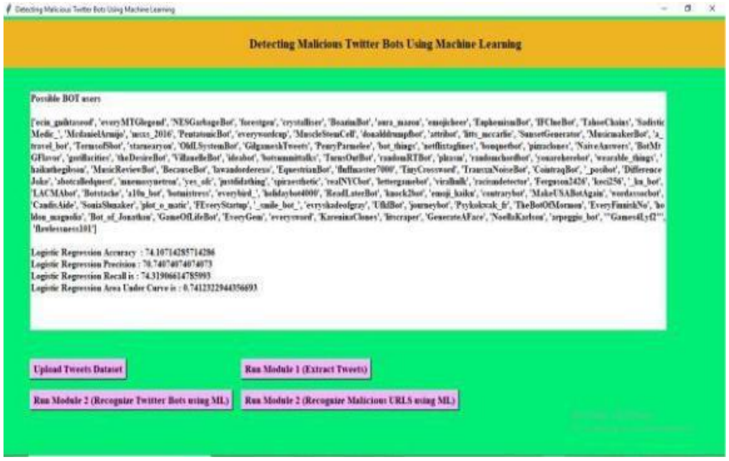
Fig.3: Uploading dataset


Detecting Malicious Twitter Bots Using Machine Learning
Detecting Malicious Twitter Bots Using Machine Learning
Upload Tweets Dataset
Run Module 1 (Extract Tweets)
Run Module 2 (Recognize Twitter Bots using ML)
Run Module 2 (Recognize Malicious URLS using ML)

```
            id             id_str     screen_name ... has_extended_profile           name bot
0    8.160000e+17  "815745789754417152"  "HoustonPokeMap" ...          False  "Houston PokŽ Alert"  1
1    4.843621e+09          4843621225        kernyeahx ...          False        Keri Nelson  1
2    4.303727e+09          4303727112  mattlieberisbot ...          False  Matt Lieber Is Bot  1
3    3.063139e+09          3063139353        sc_papers ...          False  single cell papers  1
4    2.955142e+09          2955142070     lucarivera16 ...          False       lucarivera16  1
...           ...                 ...              ... ...            ...                ... ...
2792 2.362091e+08           236209129    leslieberland ...           True     Leslie Berland  0
2793 2.545039e+08           254503881     SamuelsPants ...          False  Samuel Littlefield  0
2794 1.548544e+07            15485441       jimmyfallon ...          False        jimmy fallon  0
2795 7.513869e+08           751386872   LocalAlligator ...           True    Tammas Foxdene  0
2796 1.992936e+07            19929362     RepJoeBarton ...           True         Joe Barton  0

[2797 rows x 20 columns]
```

Fig.5: Displaying tweets



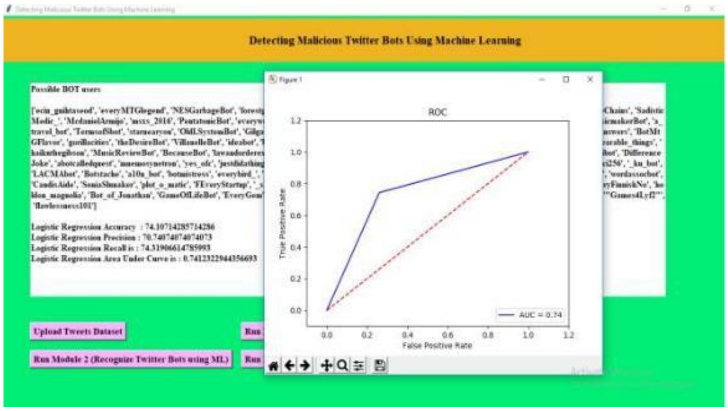
Detecting Malicious Twitter Bots Using Machine Learning
Detecting Malicious Twitter Bots Using Machine Learning
Upload Tweets Dataset
Run Module 1 (Extract Tweets)
Run Module 2 (Recognize Twitter Bots using ML)
Run Module 2 (Recognize Malicious URLS using ML)

```
Possible BOT users

['ecin_guiltasoed', 'everyMTGlegend', 'NESGarbageBot', 'forestgen', 'crystalliser', 'BoarinBot', 'aura_mazon', 'emojicheer', 'EuphemismBot', 'IFCheBot', 'TahoeChains', 'Sadistic
Medic_', 'McdanielArmijo', 'mxxs_2016', 'PentatonicBot', 'everywordcup', 'MuscleStemCell', 'donaldtrumpfbot', 'attribot', 'litts_mccarlie', 'SunsetGenerator', 'MusicmakerBot', 'a_
travel_bot', 'TermsofShot', 'starnearyou', 'OldLSystemBot', 'GilgameshTweets', 'PeeryParmelee', 'bot_things', 'netflixtaglines', 'bouquetbot', 'pizzaclones', 'NaiveAnswers', 'BotMr
GFlavor', 'gorillacities', 'theDesireBot', 'VillanelleBot', 'ideabot', 'botsummittalks', 'TurnsOutBot', 'randomRTBot', 'plzcon', 'randomchordbot', 'youareherebot', 'wearable_things', '
haikuthegibson', 'MusicReviewBot', 'BecauseBot', 'lawandorderess', 'EquestrianBot', 'fluffmaster7000', 'TinyCrossword', 'TranxxaNoiseBot', 'CointraqBot', '_posibot', 'Difference
Joke', 'abotcalledquest', 'mnemosynetron', 'yes_ofc', 'justdidathing', 'spiraesthetic', 'realNYCbot', 'lettergamebot', 'viralhulk', 'racismdetector', 'Ferguson2424', 'keci256', '_kn_bot',
'LACMAbot', 'Botstache', 'a10n_bot', 'botmistress', 'everybird_', 'holidaybot4000', 'ReadLaterBot', 'knock2bot', 'emoji_haiku', 'contrarybot', 'MakeUSABotAgain', 'wordassocbot',
'CandisAide', 'SoniaShmaker', 'plot_o_matic', 'FEveryStartup', '_smile_bot_', 'evryshadeofgray', 'UfdBot', 'journeybot', 'Psykokwak_fr', 'TheBotOfMormon', 'EveryFinnishNo', 'ho
ldon_magnolia', 'Bot_of_Jonathan', 'GameOfLifeBot', 'EveryGem', 'everyword', 'KareninaClones', 'litscraper', 'GenerateAFace', 'NoellaKarlson', 'arpeggio_bot', '"Games4Lyf2"',
 'flawlessness101']

Logistic Regression Accuracy  : 74.10714285714286
Logistic Regression Precision : 70.74074074074073
Logistic Regression Recall is : 74.31906614785993
Logistic Regression Area Under Curve is : 0.7412322944356693
```

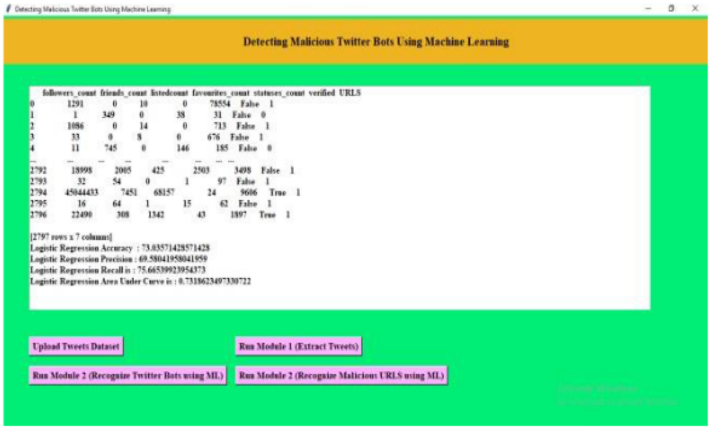
Fig.6: Possible bot users



```
Possible BOT users

Logistic Regression Accuracy  : 74.10714285714286
Logistic Regression Precision : 70.74074074074073
Logistic Regression Recall is : 74.31906614785993
Logistic Regression Area Under Curve is : 0.7412322944356693
```

Fig.7: ROC graph

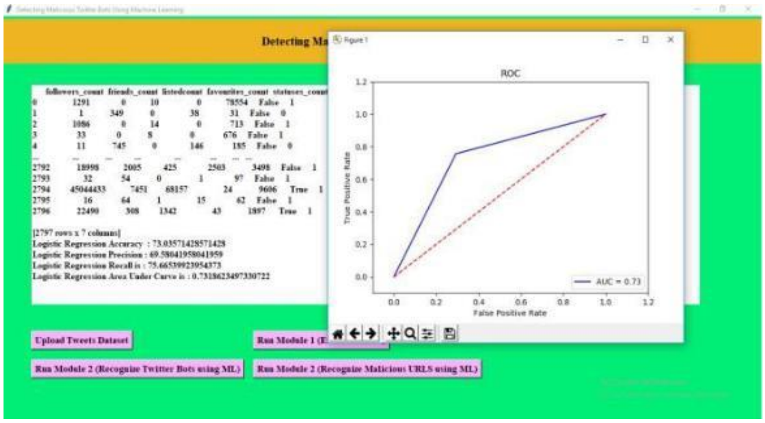


Detecting Malicious Twitter Bots Using Machine Learning
Detecting Malicious Twitter Bots Using Machine Learning
Upload Tweets Dataset
Run Module 1 (Extract Tweets)
Run Module 2 (Recognize Twitter Bots using ML)
Run Module 2 (Recognize Malicious URLS using ML)

```
   followers_count  friends_count  listedcount  favourites_count  statuses_count  verified  URLS
0           1291          0          10             0          78554   False   1
1              1        349           0            38             31   False   0
2           1086          0          14             0            713   False   1
3             33          0           8             0            676   False   1
4             11        745           0           146            185   False   0
...          ...        ...         ...           ...            ...     ...  ...
2792       18998       2005         425          2503           3498   False   1
2793          32         54           0             1             97   False   1
2794    45044433       7451       68157            24           9606    True   1
2795          16         64           1            15             62   False   1
2796       22490        308        1342            43           1897    True   1

[2797 rows x 7 columns]
Logistic Regression Accuracy  : 73.03571428571428
Logistic Regression Precision : 69.58041958041959
Logistic Regression Recall is : 75.66539923954373
Logistic Regression Area Under Curve is : 0.7318623497330722
```

Fig.8: Malicious bot users



```
     followers_count  friends_count  listedcount  favourites_count  statuses_count
0         1291        0       10          0        78554   False   1
1            1      349        0         38           31   False   0
2         1086        0       14          0          713   False   1
3           33        0        8          0          676   False   1
4           11      745        0        146          185   False   0
...        ...      ...      ...        ...          ...     ...  ...
2792     18998     2005      425       2503         3498   False   1
2793        32       54        0          1           97   False   1
2794  45044433     7451    68157         24         9606    True   1
2795        16       64        1         15           62   False   1
2796     22490      308     1342         43         1897    True   1

[2797 rows x 7 columns]
Logistic Regression Accuracy  : 73.03571428571428
Logistic Regression Precision : 69.58041958041959
Logistic Regression Recall is : 75.66539923954373
Logistic Regression Area Under Curve is : 0.7318623497330722
```

Fig.11: ROC graph

V. CONCLUSION

In our study, we proposed an algorithm to detect Twitter bots. Logistic regression was used to detect BOTS from tweets. It had a part in bringing down cybercrime. With an accuracy of around 96.7% for train data and 96.65% for test data, the bag about words approach was shown to be the best learning model when compared to decision trees, random forests, and multinomial Navie Bayes. Thus, Twitter bots were effectively discovered using real-time data and our machine learning algorithms.

FUTURE SCOPE

Real-time data may be made available in future implementations, enabling Twitter to include this feature into their service. It may also be incorporated with any other social networking programme on the market. & in this project, every detection is done manually using a dataset that we supplied. But in the future, I may improve the project so the model can use the dataset required for bot identification independently.

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 Hadiah 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.