Social Network Bot Detection using Machine Learning Algorithms

Mrs. J. Lakshmi[1], MVPNL. Thanuja[2], N. Vasudha[3], P. Praveena[4], Shaik Sohil[5]

[1]Asst. Prof, Dept of CSE, Tirumala Engineering College, Narasaraopet, Guntur, A.P., India

[2,3,4,5]B. Tech Students, Dept of CSE, Tirumala Engineering College, Narasaraopet, Guntur, A.P., India

ABSTRACT - Twitter is a popular social networking platform that allows users to share their opinions on a variety of topics such as politics, sports, the stock market, and entertainment. It is one of the quickest ways to transmit information. It has a major impact on people's perspectives. As a result, tweets must be submitted from real users rather than Twitter bots. Spam tweets are sent by a Twitter bot. As a result, identifying bots aids in the identification of spam messages. Using machine learning algorithms, this paper suggests a method for detecting Twitter bots. Decision tree, Multinomial Nave Bayes, Random Forest, and Bag of Words are all compared.

Keywords: Social Network, Twitter, Bot

I. INTRODUCTION

Twitter is one of the fastest growing social networking platforms. It enables users to post news, express their views, and debate current events. Users may follow others who share their interests or views. Users will send tweets to their followers in real time. Re-tweeting allows the content to reach a wider audience. During live events such as athletics or award ceremonies, the number of tweets increases on its own. Twitter can be reached via smartphones as well as computers. Paid campaigns will be carried out, resulting in significant revenue generation as well as increased retail revenues. Twitter provides students with additional knowledge about the subjects covered in class.

The post that is shared with the followers is referred to as a tweet. The tweet should be brief and no more than 140 characters long. The hashtag (#) is used to search and follow a specific thread. When a hashtag becomes famous, it is referred to as a trending subject. Twitter connections are bidirectional; a person may have both friends and followers. When you follow anyone on Twitter, you will be able to read all of their messages if the account is public; but, this does not guarantee that he or she will be able to see your tweets. If you follow anyone back, he or she will be able to see your tweets.

Users receive a large number of tweets, some of which are created by bots. Bot detection is needed to recognise false users and protect legitimate users from disinformation and malicious intent. A Twitter bot is programme that automatically sends messages to people. Bots are programmed to perform tasks such as spamming. Twitter bots' malicious intent is to: 1) spread lies and fake news. 2) To smear someone's reputation. 3) False messages are made with the intent of stealing credentials. 4) Users are misdirected to fraudulent websites. 5) To influence the popularity of a person or community by changing their opinions.

We're using a Kaggle dataset. It includes attributes such as the number of followers, contacts, location, screen name (used for online communication), validated (if the user is authenticated), favourite (used for liked tweets), url, id, definition, and mentioned count. The spearman correlation coefficient is used to extract features. The data collection has been learned to detect bots. We are putting Decision Tree, Multinomial Nave Bayes, Random Forest, and Bag of Words into action. To measure real-time results, the algorithm with the highest accuracy is used.

II. RELATED WORK

Machine learning models based on programmed features are used to identify fake identities produced by humans or bots. It was investigated if readily available and engineered features used for the effective identification of fake identities created by bots or machines using machine learning algorithms could also be used to detect fake identities created by humans. A dataset of features with labels classifying each row or outcome is needed for supervised machine learning algorithms. Thus, features are the feedback that supervised machine learning models use to forecast an outcome. These attributes may be attributes discovered by APIs that represent a particular piece of knowledge about an SMP account, such as the number of friends. The predictive findings from the qualified machine learning models were just 49.75 percent accurate. Without using behavioural data, the machine learning models were learned to use engineered features [1].

Content polluters, or bots who hijack a debate for political or advertisement reasons, are a well-known issue in event forecasting, election forecasting, and separating real news from false news in social networking results. This type of bot is especially difficult to identify. Content polluters are bots who threaten to derail a legitimate dialogue by hijacking it for political or commercial gain. In real time, methods were created to detect social bots in data using only partial knowledge about the user and their tweet history. They looked at two aspects of tweets: temporal detail and message diversity. It was discovered that content polluters in this dataset often coordinated their messages. The existence of bot accounts can be inferred by analysing temporal patterns. It was also discovered that bots used a limited number of URLs in their tweets [2].

Twitter users have begun to purchase bogus followers for their accounts. This will result in Twitter spam. According to an account, 13000 bought false followers and 5386 legitimate followers were manually checked. Then, a number of characteristics that differentiate between false and legitimate followers were found. These were used as attributes in machine learning algorithms to determine whether users were false or real. They provided the cumulative distribution function (CDF) for the six attributes to check that these attributes are still useful in distinguishing between false followers accounts and legitimate user accounts [3]. Bot detection is required to identify fake or malicious users and to protect genuine users from misinformation and malicious intent. Using statistic derivation, twelve features are generated that are available in the bot repository dataset, such as followers count, friends count, and so on. Other functions, such as the number of hash tags per tweet, the number of favourites per tweet, and the number of urls per tweet, are determined by aggregating across users. Logical regression, neural networks, and gradient-boosted algorithms are also used. They discovered that comparing the efficiency of these three methods gradient boosted has high accuracy in classifying users as bot or human on Twitter [4].

There are three kinds of apps: sybils, trusting users, and truthful users. An adversary's various accounts are referred to as sybil accounts. The honest and Sybil regions are sparsely linked in this area, and Sybils have a limited number of links to honest users. Sybil groups build a false trustworthy view on truthful users of the Online Social Network due to their extensive interactions. A research on profiling humans and bots revealed differences in tweet quality, tweeting activity, and account properties such as external URL ratio [5].

The associated Twitter accounts were distinguished utilizing cross corresponded exercises and no marked information not at all like the current bot identification strategies. This strategy is 94% precise and identify bots effectively [6]. Studies had shown that the greater part of the spam messages were naturally delivered by bots. Subsequently bot spammer location decreases the spam messages. Time level entropy and tweet likeness were utilized as measures for spammer discovery. Accuracy, review and f-proportion of this strategy came about in 85%,94% and 90% respectively[7]. Twitter bots that is stage or point channel contain 9% of the tweets. For each record number of tweets, adherents, followees and date of the first and last tweet were distinguished. Normal tweets each day was determined to look at the normal tweeting movement. Bot or not score on a scale from 0-100% demonstrates likelihood of twitter record to be a human or social bot. Bot or not thinks about appearance of tweets, re-tweets and notices, tweet substance and opinions. 84% of the 51 records were stage takes care of, subject feeds and specific records seemed multiple times. Stage and subject feeds delivered 4.6 and 7.1 tweets per account each day. Particular records tweeted a lot lesser than mechanized records that is 2.2 tweets each day [8].

III. PROPOSED ARCHITECTURE

The block diagram of our system is shown in figure 1 and 2.

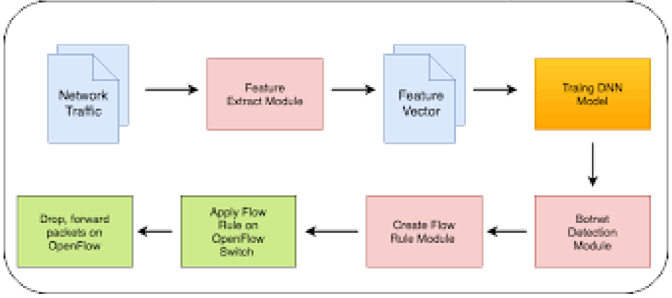

Figure 1. Proposed Architecture

There are several attributes in the train info. The Spearman correlation approach is used to obtain the appropriate functions. Three learning models are created: the Nave Bayes algorithm, the Decision Tree, the Random Forest, and the Bag of Words. Figure 1 depicts the application of the best learning model on real-time results. Using pandas, the data is preprocessed and null values are deleted (tool for preprocessing). The dataset has been trained, and the research dataset is Twitter real-time results. The result is either a 0 or a 1.

IV. RESULTS AND OBSERVATION

The true positive rate is plotted against the false positive rate to differentiate between groups in ROC curves. The train dataset is divided into 70% train data and 30% test data. True positives are those that accurately identify true ideals. A false positive is one that incorrectly labels true values as false. As seen in Figure 2, the precision for train data using the Decision Tree algorithm is 88.70% and for test data is 87.85%.

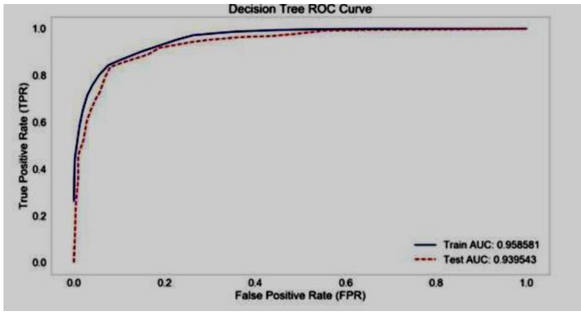

Figure 2. Results observed

The accuracy for train data using Multinomial Naïve Bayes algorithm is 67.69% and for test data is 69.76% as shown in the figure 3.

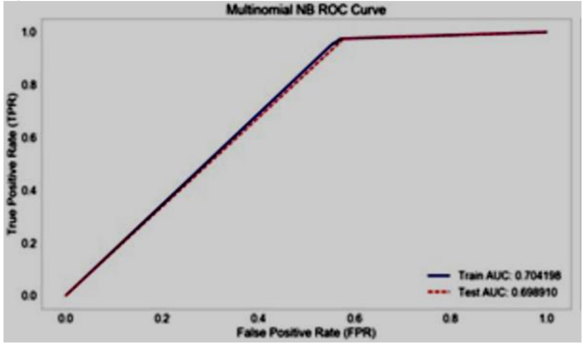

Figure 3. Results observed

V. CONCLUSION

In our article, we suggested a method for detecting Twitter bots. In relation to Decision Tree, Multinomial Nave Bayes, and Random Forest, the Bag of Words algorithm was found to be the best learning model, with an accuracy of 96.7 percent for train data and 96.65 percent for test data. As a result, the Bag of Words algorithm was used on real-time results, and Twitter bots were successfully detected.

VI. REFERENCES

[1] Van Der Walt, Estée, and Jan Eloff. "Using machine learning to detect fake identities: bots vs humans." *IEEE Access* 6 (2018): 6540-6549.

[2] Sever Nasim, Mehwish, Andrew Nguyen, Nick Lothian, Robert Cope, and Lewis Mitchell. "Real-time detection of content polluters in partially observable Twitter networks." *arXiv preprint arXiv:1804.01235* (2018).

[3] Khalil, Ashraf, Hassan Hajjdiab, and Nabeel Al-Qirim. "Detecting Fake Followers in Twitter: A Machine Learning Approach." *International Journal of Machine Learning and Computing* 7,no.6(2017).

[4] Wetstone, Jessica and Sahil R. Nayyar. “I Spot a Bot : Building a binary classifier to detect bots on Twitter.” (2017).

[5] Karataş, Arzum, and Serap Şahin. "A Review on Social Bot Detection Techniques and Research Directions." In *Proc. Int. Security and Cryptology Conference Turkey*, pp. 156-161. 2017.

[6] Chavoshi, Nikan, Hossein Hamooni, and 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, and Reddy Alexandro. "Bot spammer detection in Twitter using tweet similarity and 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, and Vincent Larivière. "Tweets as impact indicators: Examining the implications of automated “bot” accounts on T witter." *Journal of the Association for Information Science and Technology* 67, no. 1 (2016): 232-238.