IJIRCCE

e-ISSN: 2320-9801 | p-ISSN: 2320-9798

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

Volume 12, Issue 4, April 2024

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

Impact Factor: 8.379

International Journal of Innovative Research in Computer and Communication Engineering

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 | Monthly Peer Reviewed & Referred Journal |

|| Volume 12, Issue 4, April 2024 ||

| DOI: 10.15680/IJIRCCE.2024.1204163 |

Fake Identity Detection among Humans Vs Bot Using Machine Learning

Dr.N.NAVEENKUMAR, KESAVARTHINI V, KOWSALYA A, KEERTHANA P

Associate Professor, Department of CSE, Muthayammal Engineering College (Autonomous), Rasipuram, Tamil Nadu, India

Department of CSE, Muthayammal Engineering College (Autonomous), Rasipuram, Tamil Nadu, India

Department of CSE, Muthayammal Engineering College (Autonomous), Rasipuram, Tamil Nadu, India

Department of CSE, Muthayammal Engineering College (Autonomous), Rasipuram, Tamil Nadu, India

ABSTRACT: Malicious (spam) social bots generate and spread fake tweets and automate their social relationships by pretending like a follower and by creating multiple fake accounts with malicious activities. Furthermore, malicious social bots post shortened malicious URLs in the tweet in order to redirect the requests of online social networking participants to some malicious and suspicious servers. Hence, distinguishing malicious social bots from legitimate users is one of the most tasks in the Twitter network. To detect malicious or suspicious social bots, extracting URL-based features that include frequency of shared URLs, DNS fluxiness feature, network features, link popularity features and spam content presents in URLrequires less amount of time comparatively with social graph-based features (which rely on the social interactions of users). Moreover, malicious social bots cannot quickly manipulate URL redirection chains. In this, a learning automata-based malicious social bot detection (LA-MSBD) algorithm is a Machine Learning approach proposed by integrating a Naïve Bayes algorithm model with URL-based features(Fake news and Feature Extraction) for identifying trustworthy participants (users) in the Twitter network. Experimentation has been performed on 2 Twitter data sets, and the results obtained illustrate that the proposed algorithm achieves improvement in precision and detection accuracy.

KEYWORDS: Learning Automata, Fake news, Malicious Social Bots, Feature Extraction, Online social network.

INTRODUCTION

Twitter being a micro-blogging platform used by an increasing population of users of different age groups over the last decade. Generally, people post tweets and interact with other users as well. More specifically, they (users) can follow (following/friends) their favorite politicians, celebrities, athletes, entrepreneur, artists, friends and get followed by them (followers). Furthermore, Twitter generates a list of the topics being discussed day-to-day updates, that so called trending topics. Hence, users can get informed about the hot topics of discussion on a daily basis. And generally online social networks (OSNs) are increasingly used by automated accounts, well known as bots, due to their immense popularity across a wider range of user categories.It is estimated that over 15% of accounts on Twitter are automated bot accounts.A customer support chatbot is a prime example of a Twitter bot.It can help and improve the overall customer support experience by improving the response time. Following few are the most useful and amazing bots on Twitter.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 | Monthly Peer Reviewed & Referred Journal |

|| Volume 12, Issue 4, April 2024 ||

| DOI: 10.15680/IJIRCCE.2024.1204163 |

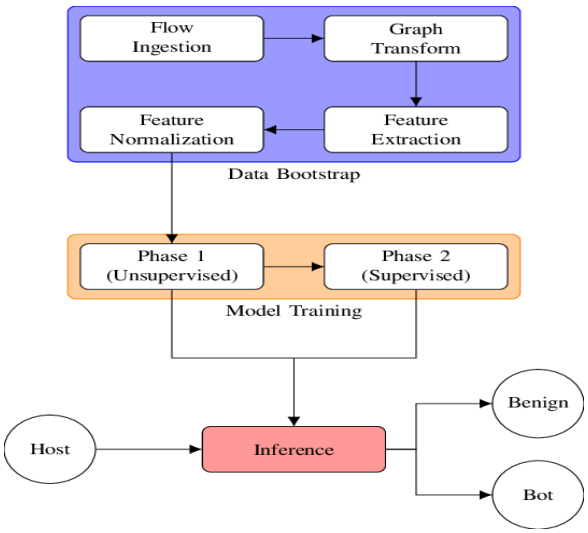

Fig 1: Bot Bootstrap

Moreover, a malicious social larva might post shortened phishing URLs within the tweet. Twitter bots are often an excellent facilitate in several distinctive ways that, there square measure cases wherever they were used unethically and illicitly. Hence, it's vital to identifying malicious social bots from legitimate users is one in all the foremost vital tasks within the Twitter network. The researchers at Indiana and North-eastern University had developed a brand new tool referred to as BOTOMETER, that tells concerning the chance of a Twitter network user being a larva. It's extremely troublesome to see AN threshold share to observe the bots however or so the score is nearer to 100%, the likelihood of the account being a larva will increase. The systems square measure being trained to acknowledge the larva behavior and analyze supported the patterns in a very dataset of over thirty,000 accounts that were initial verified by the human researchers as either bots or non-bots. Botometer a tool that "reads" over a thousand different characteristics, or "features," for each account and then assigns the account a score between 0 and 1.

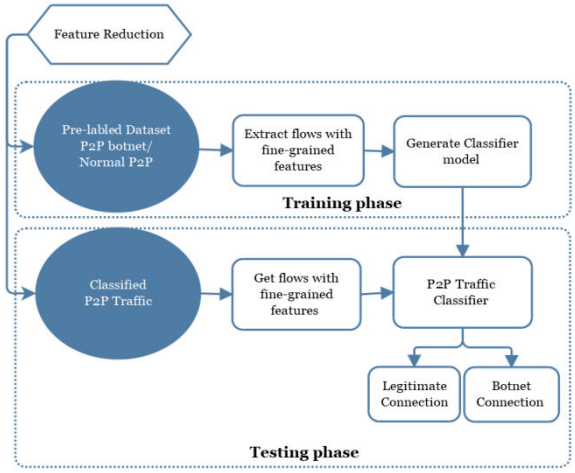

Fig 2: Bot Testing

The higher the score, the more the chances the account is automated. By the process and experiments being done one can estimate and understand the level of difficulty is too high and even time consuming to detect an twitter account is being or not and then the main task comes into role while the account if detected as bot is being purposive or subversive. Detecting an account is automated or not involves complicated steps and again detecting that automated account is malicious or legitimate is more complicated. Several techniques, including supervised, unsupervised and reinforcement learning, have been proposed to detect bots and its malicious activities in Twitter. These techniques mainly use a limited number of features extracted for identifying the automated accounts at account-level. However, there exists bots that have grown mechanisms to mimic human behavior and avoid detections. Therefore, new

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 | Monthly Peer Reviewed & Referred Journal |

|| Volume 12, Issue 4, April 2024 ||

| DOI: 10.15680/IJIRCCE.2024.1204163 |

techniques should be proposed for securing the legitimate users from the proliferation of malicious accounts in the Twitter Network.

II. LITERATURE SURVEY

In our daily lives, social media has become increasingly crucial. People naturally flock to this medium to read and s hare news, given that billions of users produce and consume information every day. Social media bots are little programmes that can be deployed on social media platforms to perform a variety of useful and destructive functions while encouraging human behaviour. Some social media bots provide helpful services like weather and sports scores. These excellent social media bots are clearly labelled as such, and those who connect with them are aware that they are bots. A huge majority of social media bots, on the other hand, are harmful bots masquerading as human users. Users lose faith in social media platforms' ability to offer accurate news as a result of these bots, since they suspect that the stories at the top of their feeds were "pushed" there by manipulative bots. Because so many individuals are using social media, malevolent users such as bots have begun to manipulate conversations in the direction that their makers desire.

These malicious bots have been used for nefarious purposes such as spreading false information about political candidates, inflating celebrities' perceived popularity, deliberately suppressing protestors' and activists' messages, illegally advertising by spamming social media with links to commercial websites, and influencing financial markets in an attempt to manipulate stock prices. Furthermore, these bots have the ability to alter the outcomes of standard social media analysis. Social media bots use a variety of attack strategies, including: Sleeper bots are bots that sleep for lengthy periods of time before waking up to unleash an attack of thousands of postings in a short period of time (perhaps as a spam attack), and then sleep again. jacking the trend - the use of top trending topics to focus on a certain audience for the purpose of targeting, An attacker employs a watering hole assault to estimate or watch which websites a company frequently visits and infects one or more of them with malware. Click farming or like farming-inflate fame or popularity on a website by like or reposting content via click farms, and hashtag hijacking- use of hashtags to focus an assault (e.g. spam, harmful links) on a specific audience using the same hashtag. In social media, bot detection is a critical duty.

Automated accounts are a problem on Twitter, a popular social networking platform. According to certain surveys, roughly 15% of Twitter accounts operate automatically or semiautomatically. The peculiarities of Twitter could be one factor that has contributed to the rise in bots. It's also worth noting that a Twitter bot is recognised as a reliable source of information. Although social networking sites have improved our social life, there are still some drawbacks. In online social networks, malicious social bots are a widespread problem. These malevolent social bots are being utilised for a variety of things, including artificially inflating a person's or movement's popularity, influencing.

Methods

We proposed as an alternative to the user-based neighborhood approach. We first consider the dimensions of the input and output of the neural network. In order to maximize the amount of training data we can feed to the network, we consider a training example to be a user profile (i.e. a row from the user-item matrix R) with one rating withheld. The loss of the network on that training example must be computed with respect to the single withheld rating. The consequence of this is that each individual rating in the training set corresponds to a training example, rather than each user. As we are interested in what is essentially a regression, we choose to use root mean squared error (RMSE) with respect to known ratings as our loss function.

Compared to the mean absolute error, root mean squared error more heavily penalizes predictions which are further off. We reason that this is good in the context of recommender system because predicting a high rating for an item the user did not enjoy significantly impacts the quality of the recommendations. On the other hand, smaller errors in prediction likely result in recommendations that are still useful—perhaps the regression is not exactly correct, but at least the highest predicted rating are likely to be relevant to the user. Data Processing is a task of converting data from a given form to a much more usable and desired form i.e. making it more meaningful and informative. Using Machine Learning algorithms, mathematical modeling and statistical knowledge, this entire process can be automated. The output of this complete process can be in any desired form like graphs, videos, charts, tables, images and many more, depending on the task we are performing and the requirements of the machine.

This might seem to be simple but when it comes to really big organizations like Twitter, Facebook, Administrative bodies like Paliament, UNESCO and health sector organizations, this entire process needs to be performed in a very structured manner.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 | Monthly Peer Reviewed & Referred Journal |

|| Volume 12, Issue 4, April 2024 ||

| DOI: 10.15680/IJIRCCE.2024.1204163 |

The most crucial step when starting with ML is to have data of good quality and accuracy. Data can be collected from any authenticated source like data.gov.in, Kaggle or UCI dataset repository.For example, while preparing for a competitive exam, students study from the best study material that they can access so that they learn the best to obtain the best results. In the same way, high-quality and accurate data will make the learning process of the model easier and better and at the time of testing, the model would yield state of the art results.

A huge amount of capital, time and resources are consumed in collecting data. Organizations or researchers have to decide what kind of data they need to execute their tasks or research. Example: Working on the Facial Expression Recognizer, needs a large number of images having a variety of human expressions. Good data ensures that the results of the model are valid and can be trusted upon.

Result Analysis

Social bots are automated social media accounts governed by software and controlled by humans at the backend. Some bots have good purposes, such as automatically posting information about news and even to provide help during emergencies. Nevertheless, bots have also been used for malicious purposes, such as for posting fake news or rumour spreading or manipulating political campaigns. There are existing mechanisms that allow for detection and removal of malicious bots automatically. However, the bot landscape changes as the bot creators use more sophisticated methods to avoid being detected. Therefore, new mechanisms for discerning between legitimate and bot accounts are much needed. Over the past few years, a few review studies contributed to the social media bot detection research by presenting a comprehensive survey on various detection methods including cutting-edge solutions like machine learning (ML)/deep learning (DL) techniques.

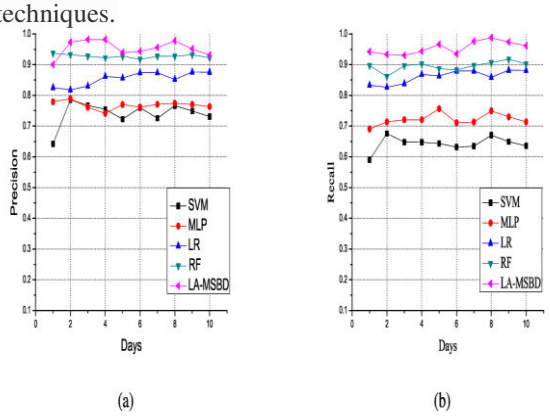

Fig 3: Result Analysis

This paper, to the best of our knowledge, is the first one to only highlight the DL techniques and compare the motivation/effectiveness of these techniques among themselves and over other methods, especially the traditional ML ones. We present here a refined taxonomy of the features used in DL studies and details about the associated pre-processing strategies required to make suitable training data for a DL model. We summarize the gaps addressed by the review papers that mentioned about DL/ML studies to provide future directions in this field. Overall, DL techniques turn out to be computation and time efficient techniques for social bot detection with better or compatible performance as traditional ML techniques.

III. CONCLUSION

The need for new, low-cost Bot detection systems is evident given the frequency of detecting malicious bots on social media sites such as Twitter. We suggested a Naive Bayes and Random Forest (RF) algorithm for detecting tweets or URLs that are potentially fraudulent or damaging to users. So far, we have downloaded and installed all of the software that is required for the planned system. The dataset was obtained from the Kaggle website, and the preparation stage was completed. The features of preprocessed data will be extracted in the next phase, and the method will be implemented, with a model saved that can be used to categories the data.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 | Monthly Peer Reviewed & Referred Journal |

|| Volume 12, Issue 4, April 2024 ||

| DOI: 10.15680/IJIRCCE.2024.1204163 |

REFERENCES

[1] P. Shi, Z. Zhang, and K.-K.-R. Choo, "Detecting malicious social bots based on clickstream sequences," IEEE Access, vol. 7, pp. 28855–28862, 2019.
[2] G. Lingam, R. R. Rout, and D. V. L. N. Somayajulu, "Adaptive deep Q-learning model for detecting social bots and influential users in online social networks," Appl. Intell., vol. 49, no. 11, pp. 3947–3964, Nov. 2019.
[3] D. Choi, J. Han, S. Chun, E. Rappos, S. Robert, and T. T. Kwon, "Bit.ly/practice: Uncovering content publishing and sharing through URL shortening services," Telematics Inform., vol. 35, no. 5, pp. 1310–1323, 2018.
[4] S. Lee and J. Kim, "Fluxing botnet command and control channels with URL shortening services," Comput. Commun., vol. 36, no. 3, pp. 320–332, Feb. 2013.
[5] S. Madisetty and M. S. Desarkar, "A neural network-based ensemble approach for spam detection in Twitter," IEEE Trans. Comput. Social Syst., vol. 5, no. 4, pp. 973–984, Dec. 2018.
[6] H. B. Kazemian and S. Ahmed, "Comparisons of machine learning techniques for detecting malicious webpages," Expert Syst. Appl., vol. 42, no. 3, pp. 1166–1177, Feb. 2015.
[7] H. Gupta, M. S. Jamal, S. Madisetty, and M. S. Desarkar, "A framework for real-time spam detection in Twitter," in Proc. 10th Int. Conf. Commun. Syst. Netw. (COMSNETS), Jan. 2018, pp. 380–383.
[8] T. Wu, S. Liu, J. Zhang, and Y. Xiang, "Twitter spam detection based on deep learning," in Proc. Australas. Comput. Sci. Week Multiconf. (ACSW), 2017, p. 3.
[9] Y. Boshmaf, I. Muslukhov, K. Beznosov, and M. Ripeanu, "Key challenges in defending against malicious socialbots," Presented at the 5th USENIX Workshop Large-Scale Exploits Emergent Threats, 2012, pp. 1–4.
[10] G. Yan, "Peri-watchdog: Hunting for hidden botnets in the periphery of online social networks," Comput. Netw., vol. 57, no. 2, pp. 540–555, Feb. 2013.
[11] D. Canali, M. Cova, G. Vigna, and C. Kruegel, "Prophiler: A fast filter for the largescale detection of malicious Web pages," in Proc. 20th Int. Conf. World Wide Web (WWW), 2011, pp. 197–206.
[12] A. K. Jain and B. B. Gupta, "A machine learning based approach for phishing detection using hyperlinks information," J. Ambient Intell. Hum. Comput., vol. 10, no. 5, pp. 2015–2028, May 2019.
[13] C. Chen, J. Zhang, X. Chen, Y. Xiang, and W. Zhou, "6 million spam tweets: A large ground truth for timely Twitter spam detection," in Proc. IEEE Int. Conf. Commun. (ICC), Jun. 2015, pp. 7065–7070.
[14] Z. Chu, S. Gianvecchio, H. Wang, and S. Jajodia, "Detecting automation of Twitter accounts: Are you a human, bot, or cyborg?" IEEE Trans. Dependable Secure Comput., vol. 9, no. 6, pp. 811–824, Nov. 2012.
[15] C. Chen, Y. Wang, J. Zhang, Y. Xiang, W. Zhou, and G. Min, "Statistical features-based realtime detection of drifted Twitter spam," IEEE Trans. Inf. Forensics Security, vol. 12, no. 4, pp. 914–925, Apr. 2017.

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

INNO SPACE SJIF Scientific Journal Impact Factor

doi crossref

निस्केयर NISCAIR

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

9940 572 462 6381 907 438 ijircce@gmail.com

www.ijircce.com

Scan to save the contact details