

Malicious Social Bots Detection in Online Social Networks with Using Ensemble Model

Dr. P. Srilatha, Kondaveti Shiva Prasad

^{1,2}Assistant Professor, Department of Computer Science and Engineering, Anurag University, TS, India

ABSTRACT:

Machine learning algorithms such as GNB (Gaussian Naïve Bayes), KNN (K-nearest neighbors), and LR (Logistic regression) are used in this paper to identify trustworthy users (participants) in OSN (online social networks) using URL-based features. Our developed model uses an ensemble of machine learning techniques for protection against malicious social bot attacks for the classification and prediction of online social networks. Data from an online social network has been used for experimentation. The proposed ensemble model accuracy is measured in terms of F-score, recall, and precision to improve accuracy and the proposed method also considers the k-fold cross-validation technique.

Keywords- LR (Logistic regression), GNB (Gaussian Naïve Bayes), KNN (K-nearest neighbors), Online social networks (OSN), k-fold cross-validation, Ensemble.

I. INTRODUCTION

Social bots are malicious software programs that appear to be actual users on OSNs. Also, social bots commit several vicious attacks, including generating fraud identities, social spam content, stealing sensitive information, and online ratings. An abbreviated URL service is used by OSN users when they want to share a tweet with others that include URLs. In addition, a malicious social bot can post criminal URLs of a tweet. When a user clicks on an abbreviated URL. The user's request will be redirected to malicious websites via intermediate URLs connected with malicious servers. After then, the attacker has access to the official participant. As a result, the OSN suffers from some disabilities (such as cybercrime). There are several suggested ways to get spam on the OSN.

Such algorithms are based on the features of tweet content and user profile. Social bots, on the other hand, can manage profile features including URL rate, hashtag ratio, as well as the number of retweets. Social bots may use touching words, icons, and the most commonly used words in tweets, by tricking each tweet's content. Users' social interactions on OSNs are difficult for malicious social bots to control. However, extracting features for social relationships is time-consuming due to the large data. So, identifying healthy bots from users' challenging tasks on the online social networks. Based on DNS information and lexical URL elements, the availability of an existing malicious URL is approaching. However, finding all social bots is difficult for finders because social bots do not deliver malicious URLs directly to tweets. Thus, it's essential to spot bad URLs (for example, malicious URLs) that social bots post on OSN.

Detection methods for malicious social bots are introduced in this article. This detection method allows classifying bot user accounts and legitimate user accounts in online social networks. To block the bot tweets and allow the legitimate tweets. To identify trustworthy participants and protect them from malicious bots,

the proposed method assembles machine learning algorithms for MSBD classification and prediction with URL-based data. The proposed model accuracy is measured in terms of F1 score, recall, and precision. To improve accuracy proposed method also considers the k-fold cross- validation technique.

The layout of the article follows the following structure:

Earlier related work is explained in Section II. Our detection approach is described in Section III. The actual implementation of the suggested approach is represented in Section IV. The proposed system's algorithms are explained in Section V. The test outcomes for different situations are represented and explained in Section VI. Lastly, Section VII summarizes our conclusion and suggests future research options.

II. RELATED WORK

We examine current strategies for detecting malicious social bots online, as well as the ensemble method.

As a result of their focus on maximizing their demands and objectives, malicious online social network members' behavior will vary from that of regular users. User behavior analysis is also one thing that helps detect the malicious social bots' accounts in OSNs in many ways. Using URL shortening and URL redirection, social bots can redirect consumers to malicious websites. It is possible to differentiate between authentic and malicious tweets based on factors like URL redirection, the number of times a URL is shared, and spam URL content. To prevent social bot intrusions [1].

Malicious social bots disrupt the network and information security by spreading misleading information and doing destructive actions. Malicious social bots on OSNs must be identified and eliminated as soon as possible.

For machine learning-based algorithms to identify social network bots, the number of real-world social bots is considerably fewer than that of real-world people, creating a significant sample imbalance issue and this result has been varied in experiments [2]. It employs machine learning to detect the primary elements of command-and-control communication (C&C) of botnets [3]. This C&C channel detection may be improved in the future by developing a better fingerprint. It may be feasible to reduce the false positive rate by considering additional information. Consider both the request and the DNS traffic when combining several requests. For the detection of HTTP and HTTPS-based C&C channels, three novel approaches have been developed. Detecting C&C channels on the network [4].

Many parts of social media have been impacted by the use of bots. Bots account for a considerable portion of Twitter's users, which has had a significant influence. These bots have been used to propagate misleading information about politicians and inflate celebrities' apparent popularity. These bots can alter the findings of popular social media analyses. Misinformation spread by malicious social bots (such as fake news) has had real-world repercussions. The detting and removal of malicious social bots from OSNs are essential [5].

Make advantage of the most recent social engineering approaches. An attempt has been made to affect the popularity of a person or a product in the target audience. Effective admission control and content-based

detection are required for better performance in terms of false positives and false negatives. Fake accounts and online social bots may be detected using a variety of approaches. Online social networks are examined from a multiagent point of view. Profiles are created and analyzed using machine learning techniques. Many strategies are discussed for detecting fraudulent accounts and associated online social bots. Online social networks from a multi-agent viewpoint have also been studied. It also goes through the machine learning approaches that may be used to create and analyze profiles [6]. Classifier models can be improved by adding social network analysis features. Plan to work on detecting which bot accounts are controlled with the same software agents. Machine learning methods may be used to identify social bots on Twitter following substantial data pre-processing and feature extraction. The study of posted tweets, personal information, and temporal patterns on Twitter user accounts yields a large number of features. Classification models using evaluation metrics including precision, accuracy, F1 Score, and recall [7].

III. PROPOSED METHOD

The primary goal of the suggested model is to identify malicious OSN accounts. Ensemble models of machine learning approaches are used in the design for classification and prediction of phishing websites from profile-based features of the participants.

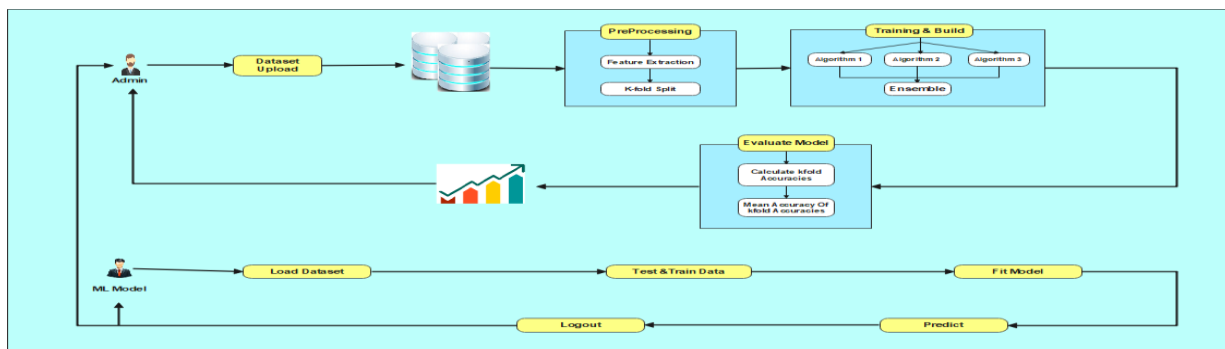


Figure 1: There are several phases in the suggested technique.

Fig. 1: This section provides an overview of the framework.

IV. METHODOLOGY

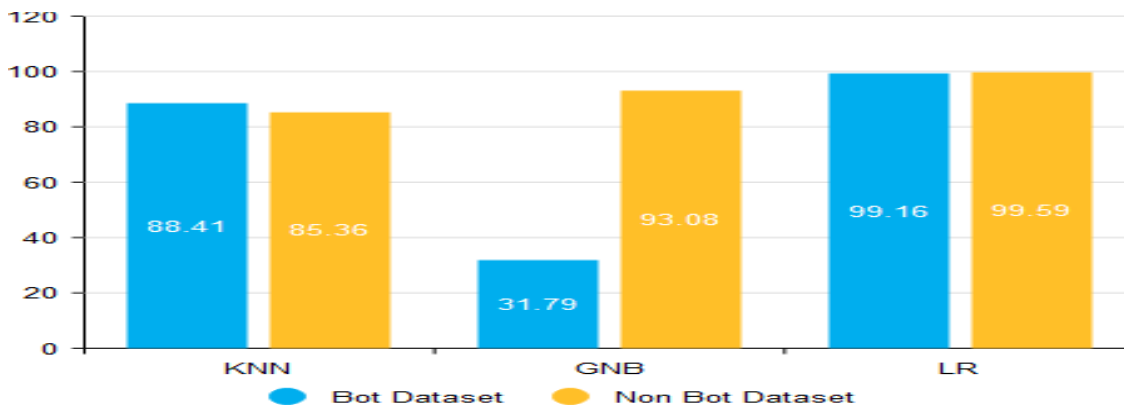
A. Dataset

Features are enumerated from a Twitter dataset of users' profiles details from online social networks (OSN). They classify the features in the picture. These attributes in the data set, from each participant's profile features, compute are ID number, URL, friends count, followers count, favorites count, listed count, default profile, statuses count, and default profile image. Figure 2 illustrates the dataset's parameters, whereas Figure 3 shows the count of malicious bots (represented by 1) and non-malicious bots (represented by 0).

Figure 2:

id	followers_count	friends_count	listed_count	favourites_count	verified	statuses_count	default_profile	default_profile_image	bot
0 8.160000e+17	1291	0	10	0	False	78554	True	False	1
1 4.843621e+09	1	349	0	38	False	31	True	False	1
2 4.303727e+09	1086	0	14	0	False	713	True	False	1
3 3.063139e+09	33	0	8	0	False	676	True	True	1
4 2.955142e+09	11	745	0	146	False	185	False	False	1

Figure 3:



B. Data pre-processing and features selection

Data must be sorted first in data pre-processing since it includes useless data. Fig.2 shows parameters of the dataset with irrelevant data, such as column Id. Attribute selection is followed by attribute subset selection, which reduces the amount of data. The dimensionality of characteristics may be lowered by studying their correlation. Correlation between features is seen in the Pearson correlation matrix shown in Figure 4. ID number, URL, Followers count, Favorites count, Listed count, verified count, default profile, Statues count, Default profile picture, Bot area are strongly associated and are picked as features for classification as shown in the graphs above.

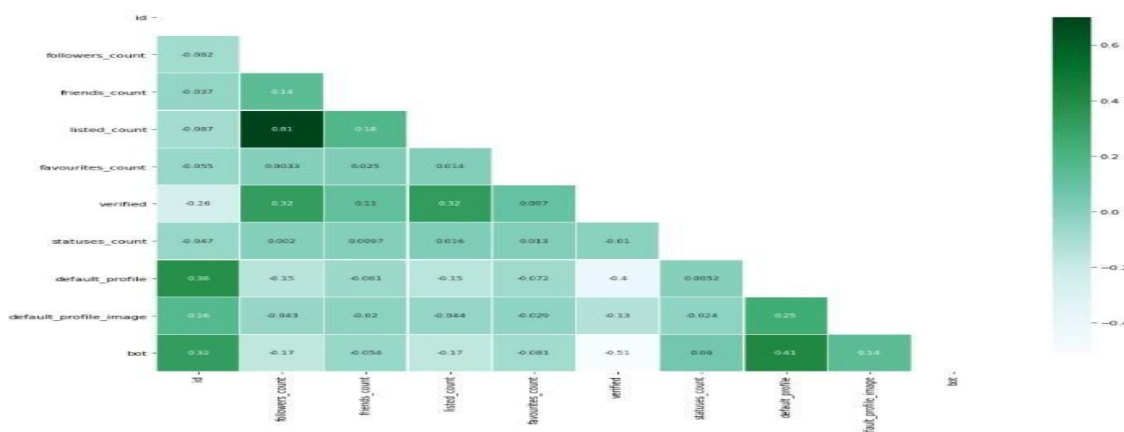


Figure 4: Pearson correlation matrix

C. Task

Classification is conducted to determine if a user's profile on an online social network belongs to a malicious social bot or an authenticated user. Only when the output variable is a category is the Classification used. Social bots and legitimate users are classified as two classes or binary classes in this classification system.

D. Machine learning algorithms

Various machine learning methods are available to address various types of issues. The algorithm is dependent on the goal, and the ensemble model is created based on this. Classification is utilized to find malicious social bots in OSN, while regression is used to measure prediction accuracy.

Predictive study of online social networks using Twitter data: a general algorithm

1. Preparing the test dataset.
2. Understanding the structure, nature, and interrelationships of data through EDA (exploratory data analysis).
3. Normalizing features, BOT Tweets in Data Sets, and other data pre-processing operations are all part of the data pre-processing process.
4. Ensemble ML Algorithms and Build a model.
5. Train & test the model.
6. Predict results.

E. Cross-validation

It is a model validation approach used to assess the performance of a machine learning ensemble model. Overfitting and underfitting may be minimized, and a sense of generalizability to a new dataset can be gained. To achieve this, divide the data into two sets: (i) training set and (ii) test set. The k-fold cross-validation approach is utilized in this study with a k value of 10. As a result, the whole data set is partitioned into ten folds and iterated ten times.

V. CLASSIFIERS

A. KNN:

Predictions are derived by scanning the whole training set for the most k comparable neighbors and summing the output variables for those k cases.

Algorithm:

- k=number of nearest neighbors is initialized.
- Determine the distance between the new instance and each of the training samples.
- Sort the distances and use the kth minimal distance to find the nearest neighbor.
- Assign the category depending on the majority category of the nearest neighbor.

B. Gaussian Naive Bayes:

It's a Naive Bayes version that uses a Gaussian normal distribution and can handle continuous data. Classification algorithms based on Naive Bayes use the Bayes theorem as their basis. It's a basic classification method with complex functionality.

Algorithm:

- We start by importing the dataset and necessary dependencies.
- Calculate Prior Probability of Classes $P(y)$.
- Calculate the Likelihood Table for all features.
- Now, Calculate the Posterior Probability for each class using the Naive Bayesian equation.

C. Logistic Regression:

A correlation between features and the probability of a certain result may be discovered using this technique. When predicting whether it will rain or not given climatic data as features, for example, the response variable has two values: Yes or No. Binomial logistic regression is the name given to this

technique. This kind of issue is known as a Multinomial logistic when the output response contains two or more possible values.

Algorithm:

- Randomly, initialize w and b .
- Calculate the predicted values of the output using the sigmoid function.
- Determine the loss by applying a loss function.
- w and b are modified such that the loss gradually diminishes to an acceptable minimal value.

VI. RESULTS

Table 2 shows the cross-validation score for $k=10$ using an ensemble model of several ML methods. Table 1 shows the accuracy based on an average of 10 iterations.

Iterations	Accuracy	Iterations	Accuracy
Kfold1	86.70	Kfold6	81.10
Kfold2	82.20	Kfold7	85.60
Kfold3	74.40	Kfold8	78.90
Kfold4	80.00	Kfold9	86.70
Kfold5	82.20	Kfold10	88.90

Table 1

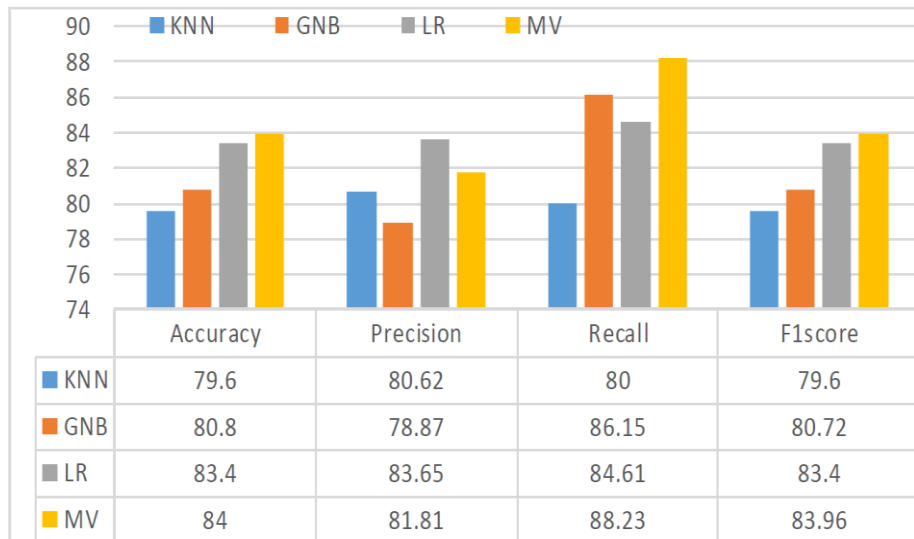
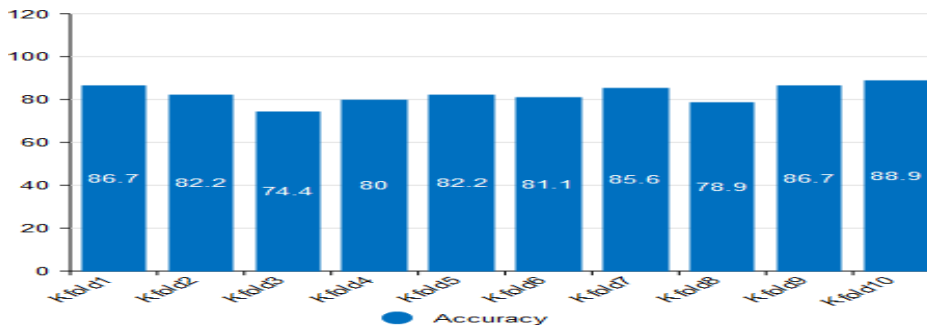


Table 2

ACCURACY MEASURES

The results of the suggested Ensemble-MSBD algorithm are compared with current machine learning models KNN, GNB, and LR based on the following metrics: Precision, Recall, and F1score.

VII. CONCLUSION AND FUTURE SCOPE

Using URL-based features from OSN participants' profiles, an ensemble-MSBD model is presented in this article. Each participant's tweets are assessed for their trustworthiness using the KNN, GNB, and LR Algorithms. A finite number of learning actions are performed using the ensemble-MSBD algorithm to update the action probability value (e.g., a participant's probability of sharing malicious URLs in their tweets).

Both attackers and defenders have been paying close attention to the fast expansion of OSNs. Researchers have suggested many social bots to tackle the threats presented by social bots, but in the future when we enter data in the malicious website or OSN account it will warn us, that it was a harmful website to close it. That may appear in the future.

VIII. REFERENCES

1. Rout RR.et.al," Detection of malicious social bots using learning automata with URL features in the Twitter network," *IEEE Trans on Comp Soci Sys*. 2020 May 14;7(4):1004-18.
2. J. Lemley.et.al, "Smart augmentation learning an optimal data augmentation strategy," *IEEE Acc*, vol. 5, pp. 5858_5869, Mar. 2017.
3. F. Tegeler.et.al,"Botfinder: Finding bots in network traffic without deep packet inspection," in Proc of the 8th intern conf on emerg netw experi and tech. ACM, 2012, pp. 349–360.
4. Kantepe M.et.al,"Preprocessing Framework for Twitter bot detection," In2017 Intern conf on com sci and engi (ubmk) 2017 Oct 5 (pp. 630-634). IEEE.
5. Shi P.et.al,"Detecting malicious social bots based on clickstream sequences," *IEEE Acc*. 2019 Feb 26; 7:28855-62.
6. Tiwari V.et.al," Analysis and detection of fake profile over social network," In2017 Intern Conf on Comp, Com and Aut (ICCCA) 2017 May 5 (pp. 175- 179). IEEE.
7. S. Lee.et.al, "Fluxing botnet command and control channels with URL shortening services," *Com. Comm*, vol. 36, no. 3, pp. 320–332, Feb. 2013.
8. Wu B.et.al," Using improved conditional generative adversarial networks to detect social bots on Twitter," *IEEE Acc*. 2020 Feb 21; 8:36664-80.
9. Lingam G.et.al," Adaptive deep Q-learning model for detecting social bots and influential users in online social networks," *Appl Inte*. 2019 Nov;49(11):3947-64.
10. Feng Y.et.al," BotFlowMon: Learning-based, content-agnostic identification of social bot traffic flows," In2019 IEEE Conf on Com and Netw Secu (CNS) 2019 Jun 10 (pp. 169-177). IEEE.
11. Morstatter F.et.al," A new approach to bot detection: Striking the balance between precision and recall," 2016 IEEE/ACM Intern Conf on Adv in Soc Net Ana and Min, ASONAM 2016 P.533-540. 7752287.

12. Reham A.et.al," Detecting social media mobile bot nets using users' activity correlation and artificial immune system," Intern Conf on Infor and Comm Sys(ICICS 2016) April 5, 2016. IEEE.
13. Kaya M.et.al," Visualization of the social bot's fingerprints," In2016 4th Intern Symp on Digi Fore and Sec (ISDFS) 2016 Apr 25 (pp. 161-166). IEEE.
14. Yadav SH.et.al," An approach for offensive text detection and prevention in Social Networks," In2015 Intern Conf on Inno in Info, Emb and Comm Sys (ICIIECS) 2015 Mar 19 (pp. 1-4). IEEE.
15. Wei CH.et.al," SimConcept: a hybrid approach for simplifying composite named entities in biomedical text," IEEE Jou of biome and hea info. 2015 Apr 13;19(4):1385-91.
16. Bhat SY.et.al," Community-based features for identifying spammers in online social networks," In2013 IEEE/ACM Intern Conf on Adv in Soc Net Ana and Min (ASONAM 2013) 2013 Aug 25 (pp. 100-107). IEEE.
17. M Van Meeteren.et.al," Mapping Communities in Lage Virtual Social Networks," - 2010 IEEE Intern Wor.