Industrial Engineering Journal
ISSN: 0970-2555
Volume : 52, Issue 3, March : 2023

SPAMMER DETECTION AND FAKE USER IDENTIFICATION ON SOCIAL NETWORKS USING ML

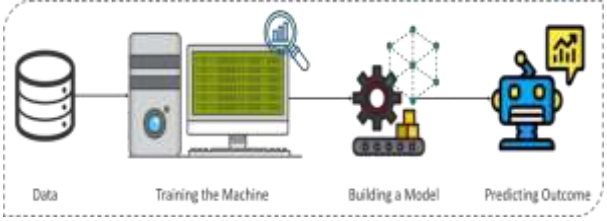
[1]Paila Anil, [2]Mylavarapu Lakshmi Bhaskar, [3]Malluvalasa Teja
Department of Computer Science and Engineering, Raghu Engineering College, Visakhapatnam
[1]19981a05b7@raghuenggcollege.in, [2]20985a0511@raghuenggcollege.in,
[3]19981a0592@raghuenggcollege.in
N.V.S.S PRABHAKAR [PHD], Assistant Professor, Department of CSE, Raghu Engineering College, Visakhapatnam, A.P., India

ABSTRACT:
The impact of social networking sites on daily life is undeniable, as millions of users worldwide engage with these platforms. Unfortunately, these platforms have also become a target for spammers to spread a large amount of irrelevant and harmful information. Twitter, in particular has become a popular platform for spammers, who use fake user identities to promote services or websites, causing disruption and consuming resources.
As a result, the detection of spammers and fake users on Twitter has become a common area of research. This paper provides a review of techniques used for detecting spammers on Twitter and presents a taxonomy that classifies these techniques based on their ability to detect fake users. The techniques are compared based on various features, including user, content, graph, structure, and time features.
The presented study aims to be a useful resource for researchers seeking to understand recent developments in Twitter spam detection. By identifying the most effective techniques for detecting spammers and fake users on Twitter, this paper provides valuable insights for researchers and practitioners seeking to improve the quality and safety of online social networks.

KEYWORDS
Classification, fake user detection, online social network, Decision Tree, spammer's identification

INTRODUCTION

It has become quite unpretentious to obtain any kind of information from any source across the world by using the Internet. The increased demand of social sites permits users to collect abundant amount of information and data about users. Huge volumes of data available on these sites also draw the attention of fake users [1]. Twitter has rapidly become an online source for acquiring real-time information about users. Twitter is an Online Social Network (OSN) where users can share anything and everything, such as news, opinions, and even their moods. Several arguments can be held over different topics, such as politics, current affairs, and important events. When a user tweets something, it is instantly conveyed to his/her followers, allowing them to outspread the received information at a much broader level [2]. With the evolution of OSNs, the need to study and analyse users' behaviours in online social platforms has intensity. Many people who do not have much information regarding the OSNs can easily be tricked by the fraudsters. There is also a demand to combat and place a control on the people who use OSNs only for advertisements and thus spam other people's accounts. Recently, the detection of spam in social networking sites attracted the attention of researchers. Spam detection is a difficult task in maintaining the security of social networks.



It is essential to recognize spams in the OSN sites to save users from various kinds of malicious attacks and to preserve their security and privacy. These hazardous manoeuvres adopted by spammers cause massive destruction of the community in the real world. Twitter spammers have various objectives, such as spreading invalid information, fake news, rumours, and spontaneous messages. Spammers achieve their malicious objectives through advertisements and several other means where they support different mailing lists and subsequently dispatch spam messages randomly to broadcast their interests. These activities cause disturbance to the original users who are known as non-spammers. In addition, it also decreases the repute of the OSN platforms. Therefore, it is essential to design a scheme to spot spammers so that corrective efforts can be taken to counter their malicious activities [3].

LITERATURE SURVEY

1. B. Erçahin, Ö. Akta³, D. Kilinç, and C. Akyol, ``Twitter fake accountdetection,'' in Proc. Int. Conf. Comput. Sci. Eng. (UBMK), Oct. 2017,pp. 388392.

Social networking sites such as Twitter and Facebook attract millions of users across the world and their interaction with social networking has affected their life. This popularity in social networking has led to different problems including the possibility of exposing incorrect information to their users through fake accounts which results to the spread of malicious content. This situation can result to a huge damage in the real world to the society.

2. F. Benevenuto, G. Magno, T. Rodrigues, and V. Almeida, ``Detectingspammers on Twitter,'' in Proc. Collaboration, Electron. Messaging, Anti-Abuse Spam Conf. (CEAS), vol. 6, Jul. 2010, p. 12.

With millions of users tweeting around the world, real time search systems and different types of mining tools are emerging to allow people tracking the repercussion of events and news on Twitter. However, although appealing as mechanisms to ease the spread of news and allow users to discuss events and post their status, these services open opportunities for new forms of spam. Trending topics, the most talked about items on Twitter at a given point in time, have been seen as an opportunity to generate traffic and revenue. Spammers post tweets containing typical words of a trending topic and URLs, usually obfuscated by URL shorteners, that lead users to completely unrelated websites. This kind of spam can contribute to de-value real time search services unless mechanisms to fight and stop spammers can be found. In this paper we consider the problem of detecting spammers on Twitter. We first collected a large dataset of Twitter that includes more than 54 million users, 1.9 billion links, and almost 1.8 billion tweets. Using tweets related to three famous trending topics from 2009, we construct a large labelled collection of users, manually classified into spammers and non-spammers. We then identify a number of characteristics related to tweet content and user social behaviour, which could potentially be used to detect spammers. We used these characteristics as attributes of machine learning process for classifying users as either spammers or non-spammers. Our strategy succeeds at detecting much of the spammers while only a small percentage of non-spammers are misclassified. Approximately 70% of spammers and 96% of non-spammers were correctly classified. Our results also highlight the most important attributes for spam detection on twitter.

3. S. Gharge, and M. Chavan, ``An integrated approach for malicious tweetsdetection using NLP,'' in Proc. Int. Conf. Inventive Commun. Comput.Technol. (ICICCT), Mar. 2017, pp. 435438.

Many previous works have focused on detection of malicious user accounts. Detecting spams or spammers on twitter has become a recent area of research in social network. However, we present a method based on two new aspects: the identification of spam-tweets without knowing previous background of the user; and the other based on analysis of language for detecting spam on twitter in such topics that are in trending at that time. Trending topics are the topics of discussion that are popular at that time. This growing micro blogging phenomenon therefore benefits spammers. Our work tries

to detect spam tweets in based on language tools. We first collected the tweets related to many trending topics, labelling them on the basis of their content which is either malicious or safe. After a labelling process we extracted a many features based on the language models using language as a tool. We also evaluate the performance and classify tweets as spam or not spam. Thus, our system can be applied for detecting spam on Twitter, focusing mainly on analysing of tweets instead of the user accounts.

PROPOSEDSYSTEM

In this article, the authors present a classification of techniques for detecting spammers on Twitter. They propose a taxonomy for identifying spammers on the platform, which is based on fake user identification. Figure 1 illustrates the taxonomy, which includes different techniques for identifying fake users.

One of the key techniques used for fake user identification on Twitter is machine learning. The authors mention several algorithms that can be used for this purpose, such as the Decision Tree algorithm, k-means clustering algorithm, and Naïve Bayes algorithm. These algorithms can be used to analyse different aspects of Twitter users' behaviour and identify patterns that are indicative of spamming activity.

Overall, the authors aim to provide a comprehensive overview of the different approaches that can be used for detecting spammers on Twitter. By presenting a taxonomy of fake user identification techniques, they hope to provide a useful resource for researchers and practitioners who are interested in combating spam on the platform. Additionally, the authors highlight the importance of identifying and countering spamming activity on social media platforms to protect users' security and privacy, and to maintain the integrity of the online community.

ADVANTAGES OF PROPOSED SYSTEM:

- This study includes the comparison of various previous methodologies proposed using different data set sand with different characteristics and accomplishments.
- Tested with real time data.
- In proposed system we use Decision Tree Algorithm. These algorithms use a smaller number of features, while still being able to correctly classify about 98% of the accounts of our training dataset.
- The results predicted by these algorithms were accurate when compared to other classification algorithms.

IMPLEMENTATION

The modules used in our system are detection, model. The model module consists of the python code that is used to build the fake user identification techniques. We made use of pandas, NumPy, matplotlib and Pylab for building the modules. And the detection modules consist of three types. Those are Decision Tree, Naive bayes and K-Means Clustering. In the first step we load the datasets and pre-processing the datasets for discretization. In the pre-processing we take the only necessary attributes from the dataset and form a new dataset to use for detecting spam profiles. In the next step the newly created dataset which came after the data pre-processing is taken as input to the three modules that are designed by using the three different algorithms. The module that is designed by using the Decision tree classifier algorithm produces the prediction with much accuracy than the other two modules that are designed by using the K-Means clustering and Naive bayes algorithms. The pickle output file is created by using flask technique for the Decision tree algorithm-based approach which is having accuracy of 98%.

The pickle output file is connected to the user interface for verifying accounts. The user interface is designed by using the HTML technology. The user interface takes the input from the end user. The end user gives the input which is the screen name of the user in the dataset. After verifying the input the output will be displayed that he is a fake user or original user.

Industrial Engineering Journal
ISSN: 0970-2555
Volume : 52, Issue 3, March : 2023

ALGORITHM

Decision Tree algorithm

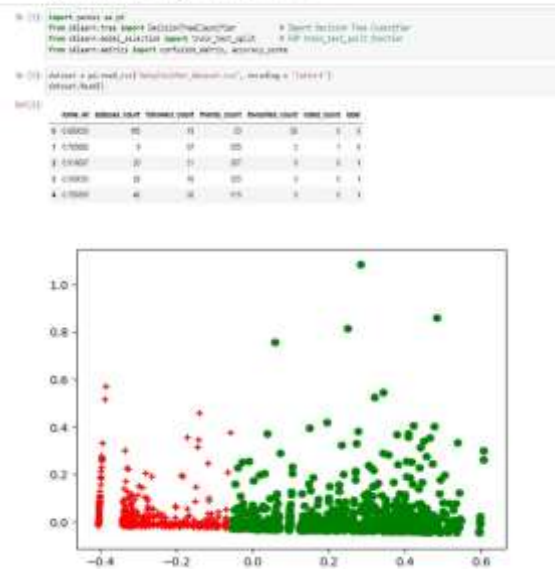
Decision trees are a popular machine learning algorithm used for both classification and regression tasks. The internal nodes of a decision tree represent a particular feature or attribute, and the branches represent the decision rules based on the values of that feature. The leaf nodes represent the outcome or prediction.

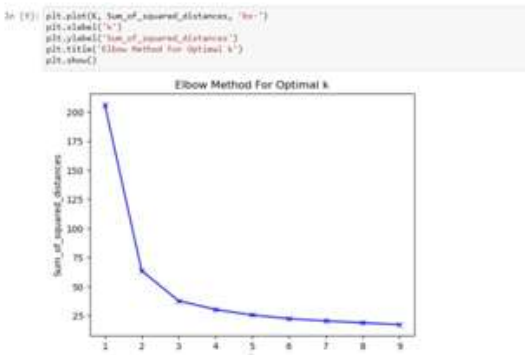
COMPARISON OF PERFORMANCE OF CLASSIFIER

Accuracy	Precision	Recall	F1-Score	Classifier
0.48	0.67	0.62	0.56	K- Means Clustering
0.98	1.0	0.97	0.98	Decision Tree
0.87	0.80	0.98	0.88	Naive Bayes

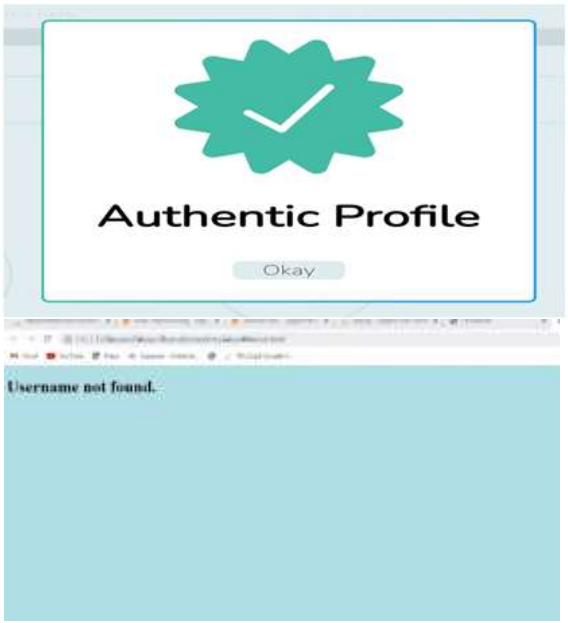
SAMPLE RESULTS

DECISION TREE CLASSIFIER





Industrial Engineering Journal
ISSN: 0970-2555
Volume : 52, Issue 3, March : 2023



CONCLUSION

Furthermore, it is also necessary to address the issue of privacy preservation while detecting spammers on social media platforms. As most of the existing spam detection techniques require access to users' data, there is a need to balance the privacy concerns with the accuracy and effectiveness of the algorithms.

Additionally, the dynamic nature of social media networks presents a challenge for spam detection techniques. The spammers often change their strategies and tactics to avoid detection, and thus, the algorithms need to be updated and adapted accordingly. Lastly, there is also a need to explore the ethical implications of spam detection techniques, especially with regard to the potential biases and discrimination that can arise from the use of such algorithms.

In conclusion, while the presented taxonomy and review of Twitter spam detection techniques provide valuable insights, there is still a need for further research in the areas of false news identification, privacy preservation, dynamic nature of social media networks, and ethical implications of spam detection techniques.

REFERENCES

B. Erçahin, Ö. Aktaş, D. Kilinç, and C. Akyol, ‘‘Twitter fake account detection,’’ in *Proc. Int. Conf. Comput. Sci. Eng. (UBMK)*, Oct. 2017, pp. 388–392.

F. Benevenuto, G. Magno, T. Rodrigues, and V. Almeida, ‘‘Detecting spammers on Twitter,’’ in *Proc. Collaboration, Electron. Messaging, AntiAbuse Spam Conf. (CEAS)*, vol. 6, Jul. 2010, p. 12.

S. Gharge, and M. Chavan, ‘‘An integrated approach for malicious tweets detection using NLP,’’ in *Proc. Int. Conf. Inventive Commun. Comput. Technol. (ICICCT)*, Mar. 2017, pp. 435–438.

T. Wu, S. Wen, Y. Xiang, and W. Zhou, ‘‘Twitter spam detection: Survey of new approaches and comparative study,’’ *Comput. Secur.*, vol. 76, pp. 265–284, Jul. 2018.

S. J. Soman, ‘‘A survey on behaviors exhibited by spammers in popular social media networks,’’ in *Proc. Int. Conf. Circuit, Power Comput. Technol. (ICCPCT)*, Mar. 2016, pp. 1–6.

A. Gupta, H. Lamba, and P. Kumaraguru, ‘‘1.00 per RT #BostonMarathon # prayforboston: Analyzing fake content on Twitter,’’ in *Proc. eCrime Researchers Summit (eCRS)*, 2013, pp. 1–12.

F. Concone, A. De Paola, G. Lo Re, and M. Morana, ‘‘Twitter analysis for real-time malware discovery,’’ in *Proc. AEIT Int. Annu. Conf.*, Sep. 2017, pp. 1–6.

N. Eshraqi, M. Jalali, and M. H. Moattar, ‘‘Detecting spam tweets in Twitter using a data stream clustering algorithm,’’ in *Proc. Int. Congr. Technol., Commun. Knowl. (ICTCK)*, Nov. 2015, pp. 347–351.
C. Chen, Y. Wang, J. Zhang, Y. Xiang, W. Zhou, and G. Min, ‘‘Statistical features-based real-time detection of drifted Twitter spam,’’ *IEEE Trans. Inf. Forensics Security*, vol. 12, no. 4, pp. 914–925, Apr. 2017.
C. Buntain and J. Golbeck, ‘‘Automatically identifying fake news in popular Twitter threads,’’ in *Proc. IEEE Int. Conf. Smart Cloud (SmartCloud)*, Nov. 2017, pp. 208–215.
C. Chen, J. Zhang, Y. Xie, Y. Xiang, W. Zhou, M. M. Hassan, A. AlElaiwi, and M. Alrubaian, ‘‘A performance evaluation of machine learning-based streaming spam tweets detection,’’ *IEEE Trans. Comput. Social Syst.*, vol. 2, no. 3, pp. 65–76, Sep. 2015.
G. Stafford and L. L. Yu, ‘‘An evaluation of the effect of spam on Twitter trending topics,’’ in *Proc. Int. Conf. Social Comput.*, Sep. 2013, pp. 373–378.
M. Mateen, M. A. Iqbal, M. Aleem, and M. A. Islam, ‘‘A hybrid approach for spam detection for Twitter,’’ in *Proc. 14th Int. Bhurban Conf. Appl. Sci. Technol. (IBCAST)*, Jan. 2017, pp. 466–471.
A. Gupta and R. Kaushal, ‘‘Improving spam detection in online social networks,’’ in *Proc. Int. Conf. Cogn. Comput. Inf. Process. (CCIP)*, Mar. 2015, pp. 1–6.
F. Fathaliani and M. Bouguessa, ‘‘A model-based approach for identifying spammers in social networks,’’ in *Proc. IEEE Int. Conf. Data Sci. Adv. Anal. (DSAA)*, Oct. 2015, pp. 1–9.