IDENTIFYING FAKE USERS & DETECTING SPAMMERS IN SOCIAL MEDIA

D Varun Prasad [1], Associate Professor**,** Department of Computer Science & Engineering, DVR & Dr. HS MIC College of Technology, Kanchikacherla, Andhra Pradesh, India
E-mail: varunprasad@mictech.ac.in
Vinay Santosh Ch[2], Assistant Professor, Department of Computer Science and Engineering, DVR & Dr. HS MIC College of Technology, Kanchikacherla, Andhra Pradesh, India
E-mail:vinaysanthosh@mictech.ac.in
Meghana M [3], UG Student, Department of Computer Science and Engineering, DVR & Dr. HS MIC College of Technology, Kanchikacherla, Andhra Pradesh, India
E-mail:meghanamorampudi123@gmail.com
Naga Sai G[4], UG Student, Department of Computer Science and Engineering, DVR & Dr. HS MIC College of Technology, Kanchikacherla, Andhra Pradesh, India, E-mail: nagasai1907@gmail.com
Harsha Vardhan M[5], UG Student, Department of Computer Science and Engineering, DVR & Dr. HS MIC College of Technology, Kanchikacherla, Andhra Pradesh, India
E-mail:matteharshavardhan@gmail.com
Jyoshna Devi V[6], UG Student, Department of Computer Science and Engineering, DVR & Dr. HS MIC College of Technology, Kanchikacherla, Andhra Pradesh, India
E-mail:devijyoshna95@gmail.com

ABSTRACT:
Online Social Networks (OSNs) are great environments for sharing ideas, following news, advertisingproducts etc., and they have been widely used by many in the world. Although these are the advantagesof social networks, it is difficult to understand whether an account in social media platform such asInstagram, Twitter, Facebook really belongs to a person or organization. Through creating fake andmalicious accounts, unwanted content can spread over the social network. Therefore, the prediction offake accounts is an important problem. In this study, we applied machine learning algorithms to thisproblemandweevaluatedperformancesofdifferentactivationfunctions.Accordingtotheexperimentalr esults, use of machine learning algorithms in detecting fake accounts yielded successful results. The useof various activation functions indifferent layers on the ANN significantly affects the results. In theliterature,otherclassificationmethodshavebeenwidelyusedfor detectingfakeaccountsandspammers on online social Network. To the best of our knowledge, there is no brief study that classifies fakeaccounts usingANNswithdifferentactivationfunctions.
KEYWORDS: Socialmedia, Artificialneuralnetwork, Spammers, Fakeprofiles.

1. INTRODUCTION:

Malicioususersproducefakeprofilestophishlogininfofromunsuspectingusers.Afakeprofilecansendfrien drequeststoseveralUserswithpublicprofiles.Thesecounterfeitprofiles baitunsuspectingusers with photos of individuals and they may misuse them for various use. Once the user accepts therequest,theownerofthephonyprofilecanspamfriend requeststoanyonethisusercould be afriend.Thefakeprofile'scontentsusuallyhavelinksthatresultinmaliciousexternalwebsitewherethereisan attack of virus to the system and when unaware curious user clicks the dangerous link can result incrashing of their systems. The effect of this may be dangerous as putting in a root kit turning the pc intoazombie.WhereasFacebookcontainsarigorousscreeningto stay thisfakeaccountsout,it solelytakesonefakeprofiletowreckthecomputersofthemany.Hence,wecameupwithasolutionbyusingmac hinelearningalgorithmswhichgavesuccessful results.To the best of our knowledge, this is an attempt to design machine learning models for the automatic classification of trend promoters. As such, our framework is generic and adaptable for tweets posted in different natural languages as it utilizes language independent features.

2. LITERATURE SURVEY:

 DOI: 10.36893.JK.2023.V13I04N16.0097-00102

ISSN: 2278-4632
Vol-13, Issue-04, No.06, April : 2023

Sybilrankwasdesignedinlate2012,towithefficiencyestablishfauxprofilesthrougharankinggraph-based system. The algorithmic rule uses a seed choice technique combined with early terminatedrandom walks to propagate trust. Its machine value is measured in O(n log n). Profile's area unit gradedconsistentwiththenumberofinteractions,tags,wallposts,andfriendsovertime.
Profilesthathaveahighrankareaunitthoughtoftoberealwithfauxprofileshavingacoffeerank withinthesystem.
First,theinformationisinJSONformat,thatisadditionalparsedtoastructuredformat(CSV) that's easier legible by machine learning techniques. These commas separated values can laterbuild the classifier additional economical. The authors tried unattended and additionally supervisedmachinelearningtechniques.Theyusedeightiethofthesamplestocoachtheclassifierandtherefor etherest to check it.
Currently, there are few experimental datasets for detecting fake reviews. Labeling data is a vasttask that extracts time and efforts. To solve the above problem, a fake review detection method based on multi-feature fusion and rolling collaborative training is proposed in the study. The novelty lies in the following two aspects: First, multiple factors such as sentiment and user behavior are integrated into a multi-level, multi feature evaluation system. They propose a method to quantify the intensity of emotions, to analyze whether the emotional tendency of the reviewer conflicts with their review behavior, so as to provide support for judging fake reviews. Second, in order to use unlabeled data to assist model learning, the authors proposed a method that uses rolling decision-making to coordinate training data so that the features extracted by the model can be dynamically updated, thereby reducing the impact of time factors on the detection performance of the classification model.

3. RELATED WORKS:

Identify Fake Reviews from The Perspective of the Review Text:
User reviews are usually short text, and fake review detection is a binary classification problem. The goal of this task is to determine whether a review is a fake review.Existing methods mainly follow the work of literature [5] and use machine learning methods to construct the classifier. Jindal and Liu [3] and others divided fake reviews into three categories: reviews involving only brands, reviews without substantial content, and untrue reviews. At that time, there were no public data sets for fake reviews and they decided if a review is fake or not by judging whether the review is a duplicate review. Yoo and Gretzel[4] and others collected hotel review data including 40 truthful review data and 42 fake review data as a dataset. In terms of linguistics,they used a standard statistical method to compare truthful reviews with fake reviews, and it was found that there were indeed differences in the expression between the two.Ott *et al.* [8] and others have built the ``Golden Standard'' in the field of fake review detection through the online crowd sourcing service provided by Amazon.
The existing systems use very fewer factors to decide whether an account is fake or not. Thefactors largely affect the way decision making occurs. When the number of factors is low, the accuracyof the decision making is reduced significantly. There is an exceptional improvement in fake accountcreation, which is unmatched by the software or application used to detect the fake account. Due to theadvancement in creation of fake account, existing methods have turned obsolete. The most commonalgorithm used by fake account detection Applications is the Naïve bias classifier. The accuracy ofexistingsystemisless comparedtoproposedsystem.
These artificially defined criteria can help us identify fake reviews. Some believe that the relationship between reviews, reviewers, and businesses can reveal the fake review activities of fake reviewers. It is proposed to use a heterogeneous review graph with three types ofnodes to capture the relationship between reviews, reviewers, and businesses reviewed by reviewers. An effective iterative algorithm for solving these three concepts based on the graphmodel has also been developed, so as to find reviewers with poor credibility and regard them as fake reviewers.
Inadequacy in Existing Research
At present, there are several problems in the related research of fake review detection:

DOI: 10.36893.JK.2023.V13I04N16.0097-00102

- Fake review detection is usually based on the classification method under the full-supervised framework. The full-supervised learning method requires a large amount of labeled data as training samples and labeled data are difficult to obtain. Manually labeling data consumes a lot of manpower and material resources, and there are inaccurate subjective labels, which will limit the progress of fully supervised learning.
- Scholars try to use unsupervised learning methods, which use unlabeled data for cluster analysis to classify through unsupervised learning, but for such more confusing detection tasks, the accuracy is not high.
- Semi-supervised learning well balances the main problems of fully supervised learning and unsupervised learning. However, in the current detection task, only the basic features such as part-of-speech or n-gram are used for modeling, and the factors such as the interaction between different features are ignored, which reduces the classification effect.
- Existing deep learning models perform well in plain text classification, but they are not effective in the field of fake review detection. The main reason is that it is difficult for fake reviews to indexed features from plain text, and it needs to be analyzed from multiple angles, such as user information, business information and other factors. Currently, a multiple dimensions method to detect fake reviews is urgently needed.

4. PROPOSED SYSTEM:

Intheproposedsystem,thesystemelaboratesaclassificationofspammerdetectiontechniques.Thesystemshowstheproposedtaxonomyforidentificationofspammerson Tw.Theproposedtaxonomyis categorized intoclasses,namely

- Fakecontent - Thefirstcategory (fakecontent)includesvarioustechniques,such asregression predictionmodel,malware alertingsystem,andLfunscheme approach.
- Fakeuser identification - The lastcategory(fakeuseridentification)isbasedondetectingfakeusersthroughhybridtechniques.

The Construction of Review Credibility Index System:

Constructing a representative feature set can effectivelyimprove the classification accuracy and generalization ability of the model. The development of the Internet makes it possible for consumers on online platforms to interact with each other. Users can share reviews and influence the purchase decisions of other consumers. The direct influence of perceived information is useful for purchase intention, and the antecedent constructs needs of information, information credibility, and information quality had a positive and significant impact on the perceived usefulness of online reviews. Information credibility is more relevant than information quality. Due to the different emphases in various studies, the diversity of review objects and platform metadata, etc., the characteristic index systems constructed in different studies also have certain differences. This article mainly examines the credibility of the review, further refines it from the two main levels of the content of the review and the behavior of the reviewer, and builds the index set from multiple perspectives.

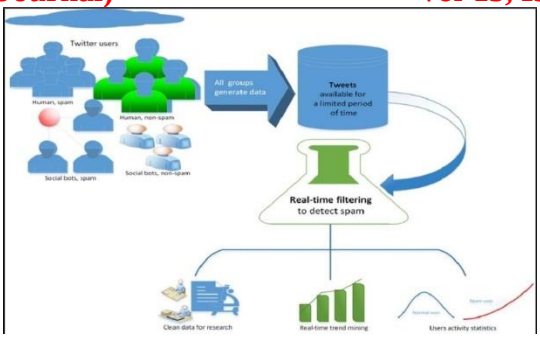
4.1 ADVANTAGESOFPROPOSEDSYSTEM

1. This study includes the comparison of various previous methodologies proposed using differentdatasetsandwithdifferentcharacteristicsandaccomplishments.
2. Testedwithrealtime data.
3. In proposed system we use Random Forest Algorithm. These algorithms use a smaller numberof features, while still being able to correctly classify about 98% of the accounts of our trainingdataset. The results predicted by these algorithms were accurate when compared to other classificationalgorithms.

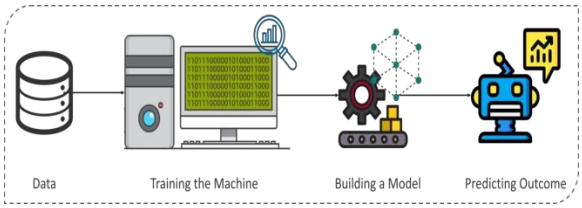
5. SYSTEMARCHITECTURE

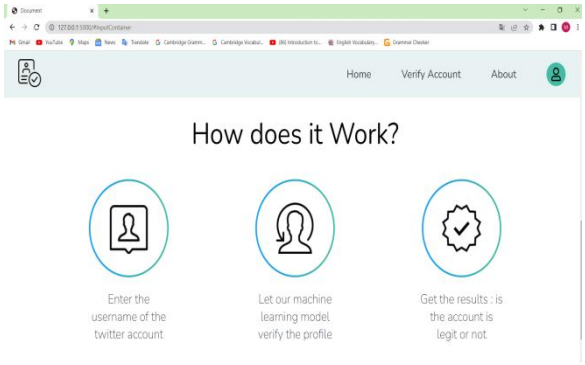
 DOI: 10.36893.JK.2023.V13I04N16.0097-00102

ISSN: 2278-4632
Vol-13, Issue-04, No.06, April : 2023



IMPLEMENTATION





The Spammer Detection and Fake User Identification on Social Networks has involving the following modules:

A. Admin Module

B. Data Collection

C. Train and Test

D. Machine Learning Technique

E. Detection of Fake User

Module Descriptions

1) Admin Module: In the first module, we develop the Online Social Networking (OSN) system module. We build up the systemwith the feature of Online Social Networking System, Twitter. Where, this module is used for admin login with their
authentication.

2) Data Collection: We will be using a Python Library called *Tweepy* to connect to the Twitter API and collect the data. We
download tweets containing certain key words, to incorporate the words or hash tags that contain relevant keyword related to
fake users.

Some of the most important fields are:

a) text, which contains the text included in the tweet.

b) created_at, which is a timestamp of when the tweet was created.

c) user, which contains information about the user that created the tweet, like the username and user id.

3) Train and Test: We present the proposed framework for metadata features are extracted from available additional informationregarding the tweets of a user, whereas content-based features aim to observe the message posting behavior of a user and thequality of the text that the user uses in posts.

DOI: 10.36893.JK.2023.V13I04N16.0097-00102

ISSN: 2278-4632
Vol-13, Issue-04, No.06, April : 2023

4) Machine Learning Technique

a) The number of features, which are associated with tweet content, and the characteristics of users are recognized for the
detection of spammers. These features are considered as the characteristics of machine learning process for categorizing users,i.e., to know whether they are spammers or not.

b) In order to recognize the approach for detecting spammers on Twitter, the labelled collection in pre-classification of fake userand legitimate user has been done. Next, those steps are taken which are needed for the construction of labeled collection andacquired various desired properties.

c) In other words, steps which are essential to be examined to develop the collection of users that can be labelled as fake user orlegitimate user. At the end, user attributes are identified based on their behavior, e.g., who they interact with and what is thefrequency of their interaction.

d) In order to confirm this instinct, features of users of the labelled collection has been checked. Two attribute sets are considered,i.e., content attributes and user behavior attributes, to differentiate one user from the other.

5) Detection of Fake User

a) In this module, we implement the collection of tweets with respect to trending topics on Twitter. After storing the tweets in aparticular file format, the tweets are subsequently analyzed.

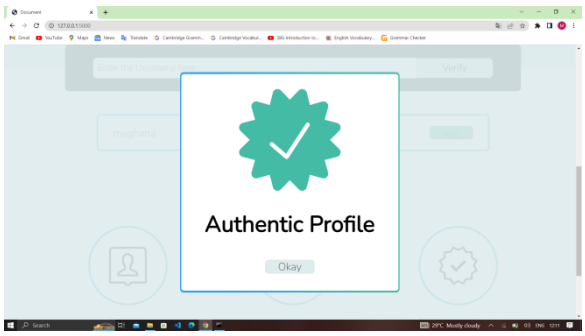
b) Labelling of fake user is performed to check through all datasets that are available to detect the malignant.

c) Feature extraction separates the characteristics construct based on the language model that uses language as a tool and helps indetermining whether the user is fake or not.

d) The classification of data set is performed by shortlisting the set of tweets that is described by the set of features provided to theclassifier to instruct the model and to acquire the knowledge for spam detection.

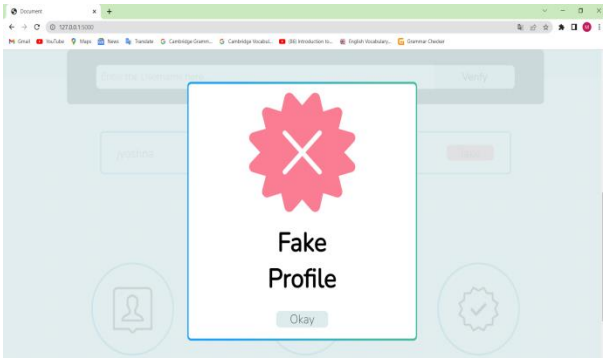
e) The fake user detection uses the classification technique to accept tweets as the input and classify the fake user and legitimateuser.

6. RESULT:



Authentic Profile:

When the username given to model, it verifies the profile and results as authentic profile if the username given is genuine.



Fake Profile:

When the username given to model, it verifies the profile and results as fake profile if the username

DOI: 10.36893.JK.2023.V13I04N16.0097-00102

ISSN: 2278-4632
Vol-13, Issue-04, No.06, April : 2023

given is not genuine.

Doesn't Exist:

When the username given to model, it verifies the profile and results as Username not found if the username given have no account.

7. CONCLUSION

We have given a framework which collects data from Twitter using Twitter API and from every tweet, we extract features that we need to feed our classifiers, that binary classification through the Random Forest is more efficient than through any other classifier. Using Decision tree, we have achieved the efficiency of 96%. In the future, we wish to classify profiles by analyzing the behavior of the user by his tweets find out a pattern and classify.Experimental results show that this method ismore effective than traditional algorithms. It uses unlabeleddata to improve the performance of the classification system,and has better classification accuracy. At the same time,the consistency of sentiment and score is analyzed, and thefeature extraction of the review is carried out through thetext representation model, and the feature fusion is combinedwith the external features of the text, which can effectively improve the classification effect of the classification model.

8. REFERENCES

Political advertising spending on Facebook between 2014 and 2018.J.R.Douceur,"Thesybilattack,"inInternationalworkshoponpeerto-peersystems.Springer,2002,pp.251–260.

Cbc.facebooksharesdroponnewsoffakeaccounts.R. Kaur and S. Singh, "A survey of data mining and social network analysis based anomaly detectiontechniques," Egyptianinformaticsjournal,vol.17,no.2,pp.199–216,2016.

L.M.PotgieterandR.Naidoo,"Factors explaininguserloyalty inasocial media-based brandcommunity," SouthAfricanJournalofInformationManagement,vol.19,no.1,pp.1–9,2017.

Quarterlyearningreports.Statista.Twitter:numberofmonthlyactiveusers2010-2018.

Y.Boshmaf,M.Ripeanu,K.Beznosov,andE.Santos-Neto,"Thwartingfakeosnaccountsbypredictingtheir victims," in Proceedings of the 8th ACM Workshop on Artificial Intelligence and Security. 2015,pp.81–89.

Facebook publishes enforcement numbers for the first time. S.-T. Sun, Y. Boshmaf, K. Hawkey, and K.Beznosov, "A billion keys, but few locks:the crisis of websingle sign-on," in Proceedings of the 2010 New Security Paradigms Workshop. ACM, 2010, pp. 61–72.

B. Arias. *How to Cover Breaking News on Twitter*. Accessed: Jan. 28, 2021. [Online]. Available: https://media.twitter.com/en/articles/best-practice/2018/how-to-cover-breaking-news-on-twitter.html

P. G. Efthimion, S. Payne, and N. Proferes, ``Supervised machine learning bot detection techniques to identify social Twitter bots," *SMU Data Sci. Rev.*, vol. 1, no. 2, p. 5, 2018.

J. Rodríguez-Ruiz, J. I. Mata-Sánchez, R. Monroy, O. Loyola-González, and A. López-Cuevas, ``A one-class classi_cation approach for bot detection on Twitter," *Comput. Secur.*, vol. 91, Apr. 2020, Art. no. 101715.

DOI: 10.36893.JK.2023.V13I04N16.0097-00102