Detection of Fake Accounts of Twitter Using SVM and NN Algorithms

Dilip Sonavane[1], Nayan Kasliwal[2], Tejas Bachhav[3], Srushti Shinde[4], Mahendra Nivangune[5]
Computer Engineering Department, Sinhgad Academy of Engineering, Pune, India[1,2,3,4,5]

Abstract- In the last years big social networks like Facebook or Twitter acknowledge that on their networks are forged and duplicate accounts. With these accounts, their creators can distribute false information, support or attack an idea, a product, or an election applicant, influencing physical network users in making a decision. They exploit the implicit belief relationships between users in order to achieve their hateful aims, for example, create hateful links within the posts/tweets.
For detecting Twitter accounts, we make use of several new features, which are more effective and robust than existing used features (e.g. number of Users/followings/followers, etc.).
We evaluated the proposed set of features by exploiting very popular machine learning classification algorithms, namely Support Vector Machine (SVM) and Neural Networks (NN). Their admiration has led to the different problems such as creation of fake accounts and spreading of fake information also creation of malicious content. Such situations may cause damage to the real-world events which are directly related to peoples, commercial entities, learning fields, etc. In this paper, we present our system build with the aim of recognizing fake users of Twitter social network.

Keywords **-** *Deep and machine learning, twitter detection, Support Vector Machine (SVM), Neural Networks (NN).*

I. INTRODUCTION

People use Twitter to share their feelings, news, events and to post their daily activities such as eating, drinking, travelling and so forth. Therefore, malicious users can check everyone's activities from their timeline and twitter becomes a place for hateful users to commit the frauds. These users which are having hostile intentions create fake accounts and spread various fake news, fake links and photos. Most of the internet users are not aware of these fake accounts; they accepted the requests and suffer in the process. Therefore, detecting fake accounts on twitter is obligatory for everyone who uses it.

In last few years, online social networks (e.g., Facebook, Twitter, and Instagram) have become one of the major platform for internet surfers to communicate with friends, express opinions, and have a talk about the events and also to cherish the memories. Twitter, a micro-blogging service launched in 2006, is one of the most popular online social network, where users post messages of around 140 characters, known as "tweet". Twitter has 330 million active users and they post nearly 500 million tweets every single day. This huge popularity attracts the attention of spammers who use twitter for hostile aims, including spreading hateful URLs within tweets, spreading rumors, sending unsolicited message to other users.

Online Social Networks (OSNs) have also attracted the interest of researchers for mining and analyzing their massive amount of data, exploring and studying user's behaviors as well as detecting their preternatural activities

Researchers have made a study to predict, analyze and explain customer's loyalty towards a social media-based online brand community, by identifying the most effective cognitive features that predict the customer's attitude.

II. EXISTING SYSTEM

In the paper [1] a review of number of data mining approaches used to detect anomalies. An uncommon direction is made to the investigation of informal community driven irregularity identification systems which are comprehensively named execution based, structure based and phantom based. Every last one from these gatherings further combines to number of procedures which are discussed in the paper. The paper has been closed with various future headings and territories of research that could be tended to and worked upon.

In paper [2] the main purpose of the detection is to secure the accounts from manipulation from fake users. We gather in excess of 69 million tweets from 5 million records. Utilizing the gathered tweets, we first direct an information investigation and find proof of Twitter pattern control. At that point, we learn at the subject level and derive the key factors that can decide if a theme starts slanting because of its prevalence, inclusion, transmission, potential inclusion, or notoriety. What we find is that with the exception of transmission, all of elements above are firmly identified with inclining. At long last, we further research the inclining control from the viewpoint of traded off and phony records and talk about countermeasures.

In paper [3], over time, social network users build trust relationships with the accounts they follow. This conviction can create for an assortment of reasons. For instance, the client may know the proprietor of the believed record face to face, or the record may be worked by an element regularly considered as dependable, for example, a general news office. The power of a record fall under the control of a digital lawbreaker, he can use this trust to facilitate his pernicious content. Here the paper shows how the comparable methods to distinguish between bargains of individual high-profile accounts can be utilized. High-profile accounts much of the time have one trademark that makes this location solid they show reliable conduct after some time. We demonstrate that our framework, was it sent, would have had the option to distinguish and counteract three true assaults against well-known organizations and news exercises. Moreover, our framework, as opposed to well-known media, would not have fallen for an organized trade-off affected by a US café network for attention reasons.

In paper [4], they proposed a new credibility analysis system for assessing information authority on Twitter to prevent the proliferation of sham or malicious information. The proposed framework comprises of four incorporated segments: a notoriety-based segment, a validity classifier motor, a user experience component, and a feature ranking algorithm. These segments operate together in an algorithmic structure to study and access the reliability of Twitter tweets and users. They also tested the performance of a system based on two different datasets from 489,330 premium twitter accounts.

The paper [5] puts forward the research discussed in this paper and applies these same engineered features to a set of fake human accounts in the hope of advancing the successful detection of fake identities created by humans on SMPs. A Diverse way to deal with contemplating validity on Twitter: they looked to demonstrate how name worth inclination influences the decisions of microblog creators. In this investigation, the creator demonstrated the relationship between name worth inclination and the number of adherents. Accordingly, it is hard to gauge the validity of a client in these systems and to check his/her posts. As online interpersonal organizations have turned out to be progressively valuable for dispersing data to more extensive crowds, tending to the previously mentioned difficulties to decide the validity of clients in OSNs requires the advancement of hearty methods for estimating client and substance believability.

We investigate the nature of spam users on Twitter with the goal to improve existing spam detection mechanisms. For detecting Twitter spammers, they have used several new features, which are more effective and robust than existing used features.

Based on the anticipated results from the machine learning models, it seems that existing features and machine learning models used to detect fake accounts are not enough to detect fake social networking accounts.

We analyzed the tergiversation idea that were used by Twitter spammers. They observed that Twitter spammer used to change their performance to evade spam detection techniques, so they suggested to design a new feature that would enhance detecting spammers and would be harder for them to evade. They had combined their new features in four machine learning classifiers and compared the implementation with other existing methods.

Drawbacks or Limitations:

Twitter has the spam problem like other social networking sites, some twitter users only tweet their products, blog or website links, some users send the spam messages or they spam you by tweeting the spam tweets, Twitter also faces the overloading problem means due to the large numbers of users and it gets crashed.

III. SYSTEM OVERVIEW

In this paper our target is to use Machine learning classification algorithms to decide the target accounts identity as real or fake, those algorithms were support vector machine, neural Network, and our newly developed algorithm, SVM-NN. The proposed algorithm (SVM-NN) uses less number of features, while its still being able to correctly classify about of the accounts of our training dataset.

Fig. 1 shows the proposed system architecture of Twitter fake account detection. The input traffic data is used for twitter dataset with some features. The training dataset contains data preprocessing which includes two steps: Feature Extraction and Machine learning technique. After usage these two are arranged in a model, which used for selecting number of features. After that apply the Support Vector machine for classifying our data and neural network use for training our model. After applying the algorithms, it predicts our model whether is the account is fake or not.

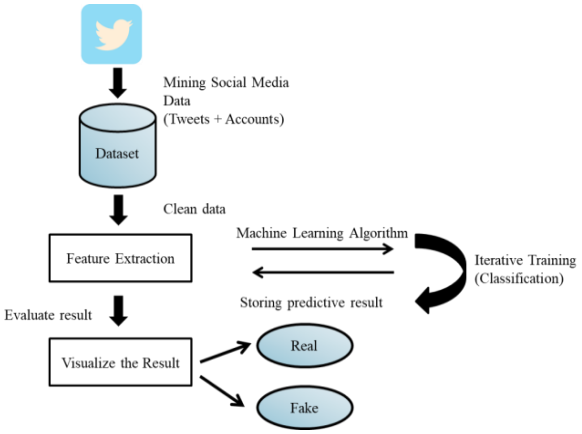


Fig. 1 Proposed System Architecture

Advantages of Proposed System:

- Due to machine learning technique, it improves accuracy of fake account detection system.
- The network or computer is constantly monitored for any invasion or attack.
- The major advantage of Support Vector Machine that classify our composite data model and predict high accuracy and NN Neural Network works to train the model with less time of work.
- Twitter's major advantage is, Twitter has limited message size, it has 140 characters per post, it can include a message or link on your website as it is free and also free for the advertisements, you don't have to face the problem with bunch of posters like the other social networking

IV. RESULT ANALYSIS

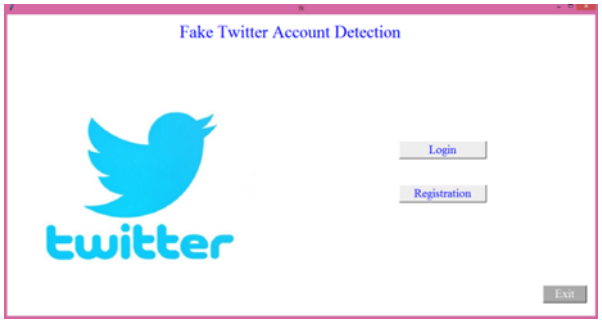


Fig.2 Home Page

Fig. 3 Registration Page

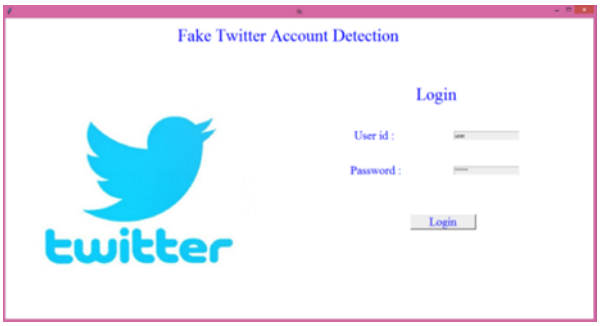


Fig. 4 Login Page

Fig. 5 Detect Fake Twitter Account

Fig. 6 View Result

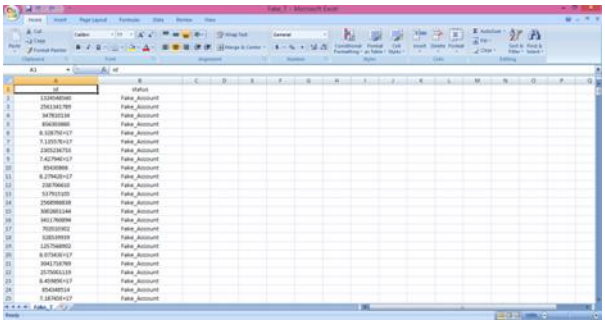
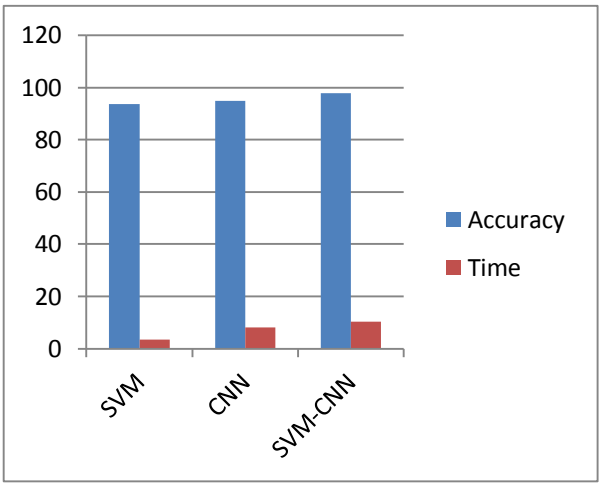


Fig. 7 View Excel Result

Comparative Analysis

Algorithm	Accuracy	Time
SVM	93.71	3.5
CNN	94.87	8.17
SVM-CNN	97.87	10.31



V. CONCLUSION

In this paper, we have maintained the highest accuracy in detecting fake accounts by different classifying algorithms. The result shows the increase of the accuracy results of five of the classification algorithms after using the suggested attributes with their corresponding heaviness. The classification algorithms are proposed to improve detecting fake accounts on social networks, where the SVM trained decision values were used to train a NN model, and SVM testing decision values were used to test the NN model.

VI. REFERENCES

[1] R.Kaur and S.Singh, "A survey of data mining and social network analysis based anomaly detection techniques", Egyptian informatics diary, vol.17, no.2, pp.1992-216, 2016.

[2] Yubao Zhang, Xin Ruan, Haining Wang, Hui Wang, and Su He "Twitter Trends Manipulation: A First Look Inside the Security of Twitter Trending" IEEE Exchanges on Data Crime scene investigation and Security (Volume: 12 , Issue: 1 , Jan. 2017)

[3] Manuel Egele, Gianluca Stringhini, Christopher Kruegel, and Giovanni Vigna "Towards Detecting Compromised Accounts on Social Networks", IEEE Exchanges on Reliable and Secure Processing (Volume: 14 , Issue: 4 , July-Aug. 1 2017)

[4] Majed Alrubaian, Muhammad Al-Qurishi, Mohammad Mehedi Hassan, and Atif Alamri, "A Credibility Analysis System for Assessing Information on Twitter", IEEE Exchanges on Reliable and Secure Processing (Volume: 15 , Issue: 4 , July-Aug. 1 2018)

[5] ESTEE VAN DER WALT and JAN ELOFF "Using Machine Learning to Detect Fake Identities: Bots vs Humans" Received December 5, 2017, acknowledged January 12, 2018, date of production January 23, 2018, date of current rendition Walk 9, 2018.

[6] Myo Myo Swe and Nyein Nyein Myo, "Fake accounts on twitter using Blacklist," in International Conference on Information System (ICIS), 2018 International Conference on. IEEE, 2018, pp. 562–566.

[7] Naman Singh, Tushar Sharma, Abha Thakral, Tanupriya Chaudhury, "Detection of Fake profile in Online Social Network using Machine Learning," in International Conference on Advances in Computing and Communication Engineering (ICACCE), 2018 International Conference on. IEEE, 2018, pp. 231-234.

[8] Ilham Aydin, Mehmet SEVİ, Mehmet Umut SALUR, "Detection of Fake Twitter Account with Machine Learning Algorithms," in International Conference on Artificial Intelligence and Data Processing (IDAP), 2018 International Conference on. IEEE, 2019.