Detection Of Fake Profile in Online Social Networks Using Machine Learning

Sonal Vinod Chaudhari
Shram Sadhana Bombay Trust's College of Engineering and Technology, Jalgaon, Maharashtra

Dipika Sanjay Mahajan
Shram Sadhana Bombay Trust's College of Engineering and Technology, Jalgaon, Maharashtra

Chaitali Anil Patil
Shram Sadhana Bombay Trust's College of Engineering and Technology, Jalgaon, Maharashtra

Priyanka Kailas Chavan
Shram Sadhana Bombay Trust's College of Engineering and Technology, Jalgaon, Maharashtra

Karishma Suresh Falak
Shram Sadhana Bombay Trust's College of Engineering and Technology, Jalgaon, Maharashtra

ABSTRACT

Social media sites are used on a regular basis in today's world and have become an integral part of our lives. It is one of the main means of communication and has become a tool for both spammers and scammers. Such social media platforms have changed drastically how we live our social life. It has become easier to make new friends, keep in touch with them, and know their updates. But many problems, such as fake profiles and online impersonation, have also grown with the rapid growth of social media. This project presents a machine learning (Neural Network) method for detecting fake accounts in social media. In this work a study is carried out of different research papers and provides a comparative study of existing work done. Fake accounts are mostly used by intruders to perform malicious activities such as personal identity harm, theft, and privacy intrusion on social me.

Keywords— *Social media, Fake accounts, Machine learning algorithms, Comprehensive Review*.

1. INTRODUCTION

These days, online social network is rising extremely fast. It is very important for marketing companies who focuses on to promote themselves by creating a base of followers and fans. Social Media such as Facebook and Instagram have become increasingly popular and vital part of todays era. The use of social media as a medium of communication, it is often
used to boost popularity and support businesses. At first, the popularity of an account is measured by some metrics such as follower count or shared contents such as the number of likes, comments, or views. Social media is a great platform for our lives, but there are diverse issues related to social media which need to be addressed. Issues related to social media, such as confidentiality, online abuse, misuse, and bullying, etc. are most commonly used by fake accounts that seems to have been created on behalf of organizations or individuals , which can damage reputation and reduce the number of likes and followers of individuals. On the other hand, fake account creation is expected to cause more damage than any other form of cybercrime.
There are various reasons for making fake accounts on social media introduces some reasons. Some reasons why people make fake accounts are:
1. Social Engineering
2. Online impersonation
3. Advertising and Campaigning
4. Privacy Intrusion etc.
Generally, all social media spammers are legal users. It is therefore a challenge to recognize them, in addition to recognizing them from legal users. Moreover, fraudsters can still use cheap automated approaches. Acquiring credibility in addition to trust and making it hard for the huge population of social media users to perceive. Social media fake account identification is a classification problem in which legal users are well recognized from fake user based on their corresponding features. Identity is an attribute that is connected to a human being, apart from him or her.

1.1 BACKGROUND

The social media networks are being used on daily basis and has become an important part of our lives. The number of people on social media platforms are increasing at a greater level for malicious use. Social networking sites in early 1997, then in 2000, could not survive much and closed very soon. As compared to those the new sites like LinkedIn, became more successful and facebook was launched in 2004 and presently it is the largest social networking site in the world. Users gained more unnecessary knowledge during surfing which are posted by fake users. Researches have observed that 20 % to 40 % accounts in online social networks like facebook, instagram, twitter, etc are fake accounts. Thus the detection of fake accounts in online social networks results into solution using frameworks.

1.2 MOTIVATION

There are many social networking sites including Twitter, Facebook, Google+, Instagram,and LinkedIn. There were 823 million people who used Facebook daily on their devices, which is an increase from the 654 million such users in the previous quarter. I have been in multiple instances wherein I read a review online for a movie or a restaurant which was not as suggested. These reviews are usually done by anonymous fake users reviewing places to promote it. This is a simple example that I have come across but there are similar issues where spam users spread false content online. This may also lead in harassment of a real user or try to retrieve account information of real users. These instances have been increasing day by day and thus require a solution. No one has come up with feasible solution to such problem. In this project we plan to give a framework with which we will be able to do the detection of fake profiles can be done so that the social life of people become secured**.**

1.3 PROBLEM STATEMENT

Social media provides easier and accessible to connect and communicate with people, how we interact the quality of interpersonal relationships. However, the quality of interpersonal is at risk. Internet is interwoven into daily life, people has made social media applications indispensable to the life. Variety usage of social media continue increases with advance technology that made the world a domestics world where people connect to anyone freely from diverse place. Interaction with strangers across worldwide has make it possible and opportunity for hacker to access individuals vital information and cyber crime happen. As a result, social media has a huge impact that could affect interpersonal relationship when
people depend strongly and allow social media control the communication.

SYSTEM OBJECTIVE

In todays online world there have been a lot of problems like fake profiles, online frauds, fake reviews, etc. To date, no one has come up with a solution to these problems. The main Objective is to detect spam users from online platform and block their account from those platforms. It may reduce many online spamming incidents in the future.

1. To define the scope of the field to detect the fake accounts on social networking sites.
2. To study the various machine learning techniques used for classification of data.
3. To study the traits for detecting the fake accounts.
4. To identify the required and proper techniques for desire results.
5. To make life of people secured.

3. LITERATURE SURVEY

Yasyn Elyusufi et al.[2020] This paper proposed an approach to the detection of a false profile on the social media site using minimal profile data. The proposed model was trained independently using the supervised learning algorithm for data sets, including fake and legitimate users. The ensemble classifier was used to make predictions more accurate. Three supervised machine learning algorithms are used in this research work. Random Forest, Decision Tree and Naive Bayes are used to identify false and genuine profiles. The result shows that the Random forest algorithm is better than the other Algorithm with a precision score of 99.64 percent.

4. PROPOSED SYSTEM

The system will consist of two parts. The First part will be for the user for login and registration. The Second part will be for checking. Firstly, the user must have to register himself/herself on site, after registration user will be able to log in and will get authority. People having the correct access authority can only operate the data.

Then for the checking window, it has the Second part where the classifier will be permitted only for taking attributes from the user and show results.

5. SYSTEM DESIGN

The process of using this framework goes like in the start with the registration of user. Profile is to be selected that need to be tested. Once profile is selected suitable attributes (i.e. features) is selected on which the classification algorithm is implemented. The attributes are then passed to the trained classifier. The classifier gets trained regularly as new training data is feed into the classifier. The classifier determines whether the profile is spam or real.

6. FIGURES

6.1 Architecture of the system

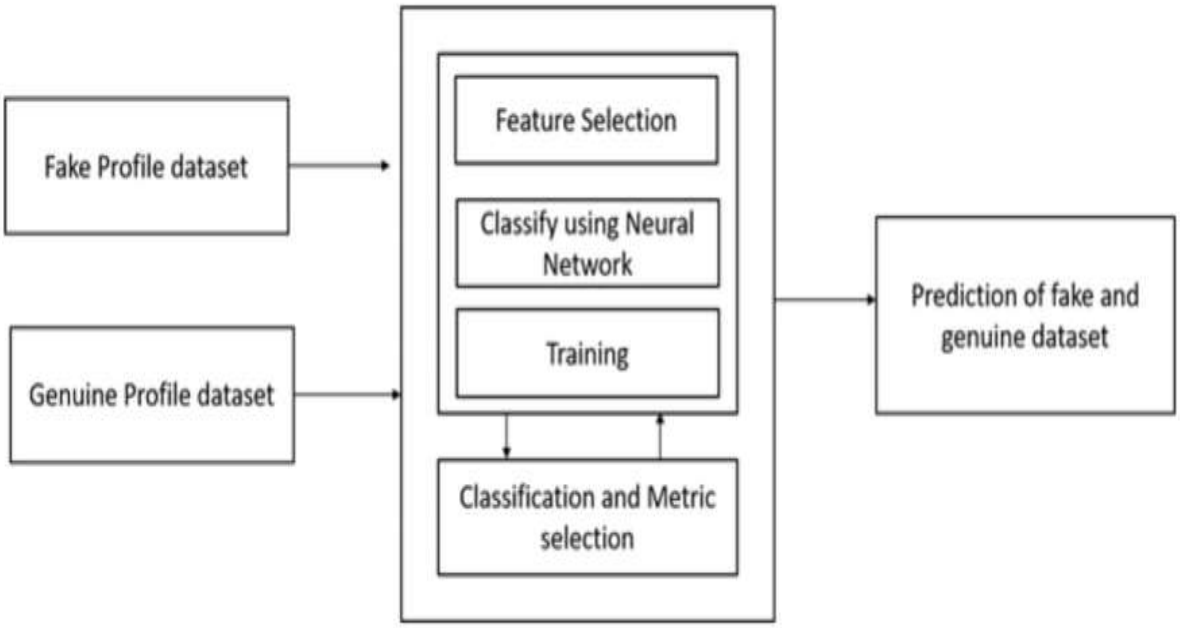

Fig. 1: Architecture of the system

6.2 Data flow diagram

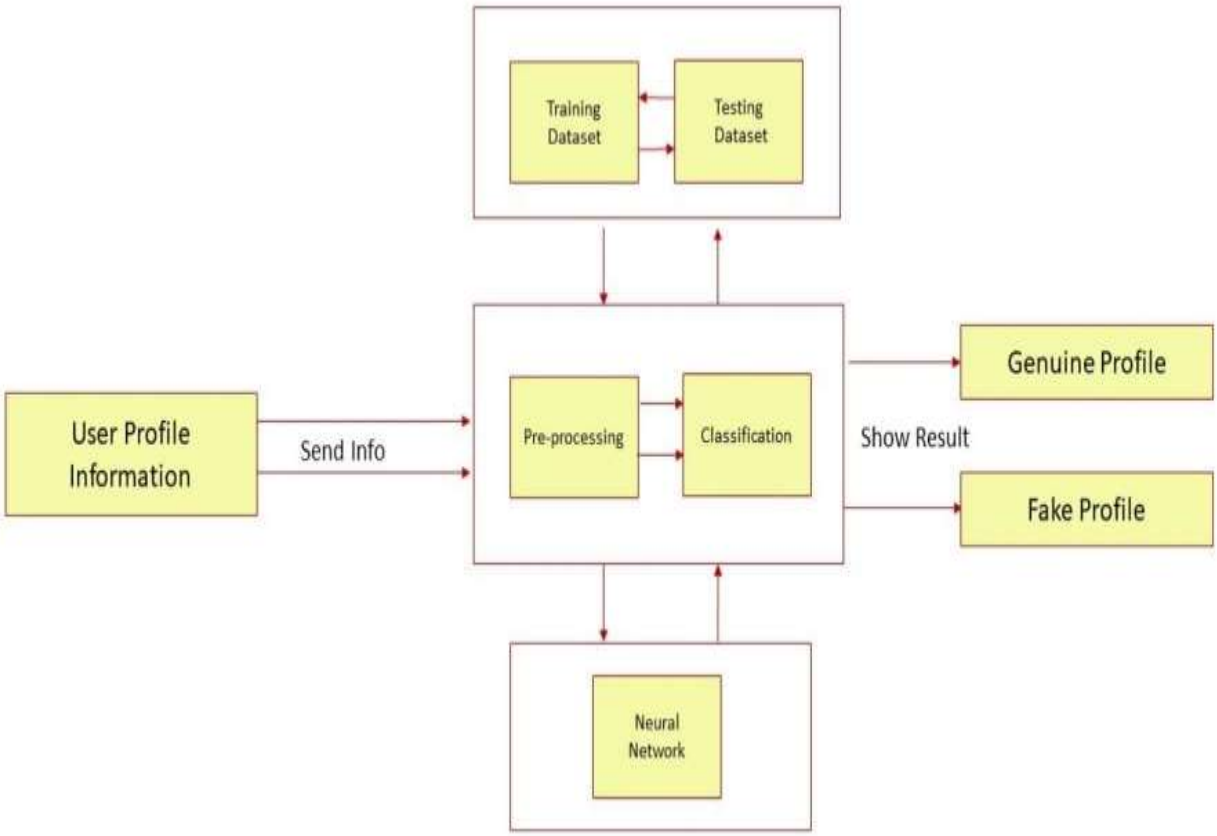

Fig. 2: Data flow diagram

7. USED ALGORITHMS

Fake Profile Detection Project uses a Convolutional Neural Networks(CNN)for Classification and Prediction.A Convolutional Neural Network (ConvNet/CNN) could also be a Deep Learning algorithm which can absorb an input image, assign importance to varied aspects/objects within the image and be ready to differentiate one from the opposite . The pre-processing required in a Convolutional Neural Networks is much lower as compared to other classification

algorithms. The architecture of a Convolutional Neural Networks is analogous to that of the connectivity pattern of Neurons within the Human Brain and was inspired by the organization of the Visual Cortex. Individual neurons answer stimuli only during a restricted region of the field of vision known as the Receptive Field.

8. RESULT

Base on the user input the classifier takes the inputs and show the results if the profile is fake or not. If prediction value is -1 then Profile is real otherwise it is Fake.Classifier is trained regularly as new data is fed into the classifier.

9. CONCLUSION AND FUTURE WORK

In this work, we proposed an approach to identify the fake profile in social network using limited profile data, about thousands of users. As we concluded in our project, we demonstrate that with limited profile data our approach can identify the fake profile with 85.64% correctly classified instances and only 14.36 % incorrectly classified instances, which is comparable to the results obtained by other existing approaches supported the larger data set and more profile information. Our research are often a motivation to figure on limited social network information and find solutions to form better decision through authentic data. Additionally, we will attempt similar approaches in other domains to seek out successful solutions to the matter where the smallest

amount amount of data is out there . In future effort will be made a new enhanced algorithm will be proposed based on these techniques to stop fake profile detection on social networks.

10.REFERENCES

[1] Boshmaf, Y., Muslukhov, I., Beznosov, K., Ripeanu, M.: The socialbot network: when bots socialize for fame and money. In: Proceedings of the 27th Annual Computer Security Applications Conference, pp. 93102. ACM (2011)

[2] Romanov, A., Semenov, A., Veijalainen, J.: Revealing fake profiles in social networks by longitudinal data analysis. In: 13th International Conference on Web Information Systems and Technologies, January 2017

[3] Song, J., Lee, S., Kim, J.: CrowdTarget: target-based detection of crowdturfing in online social networks. In: Proceedings of the 22nd ACM SIGSAC Conference on Computer and Communications Security, CCS 2015, pp. 793804. ACM, New York (2015)

[4] Nazir, A., Raza, S., Chuah, C.-N., Schipper, B.: Ghostbusting Facebook: detecting and characterizing phantom profiles in online social gaming applications. In: Proceedings of the 3rd Conference on Online Social Networks, WOSN 2010. USENIX Association, Berkeley, CA, USA, p. 1 (2010)

[5] Adikari, S., Dutta, K.: Identifying fake profiles in Linkedin. Presented at the Pacific Asia Conference on Information Systems PACIS 2014 Proceedings (2014)

[6] Stringhini, G., Kruegel, C., Vigna, G.: Detecting spammers on social networks. In: Proceedings of the 26th Annual Computer Security Applications Conference, ACSAC
2010, pp. 19 (2010)

[7] Yang, C., Harkreader, R.C., Gu, G.: Die free or live hard? Empirical evaluation and new design for fighting evolving Twitter spammers. In: Proceedings of the 14th International Conference on Recent Advances in Intrusion Detection, RAID 2011, pp. 318337. Springer, Heidelberg (2011)

[8] Elyusufi, Y., Seghiouer, H., Alimam, M.A.: Building profiles based on ontology for recommendation custom interfaces. In: International Conference on Multimedia Computing and Systems (ICMCS) Anonymous IEEE, pp. 558562 (2014)