International Journal of Scientific Modern Research and Technology
(Volume: 11, Issue: 1, Number: 4)
ISSN: 2582-8150

Fake Account Detection in Social-Media: A Comprehensive Review

Prabhakar Azad[1], Neelesh Rai[2]
[1]Research Scholar, [2]Head and Assistant Professor
Department of CSE, MITS, Bhopal, India

Abstract- These days, social networking services are ubiquitous and almost indispensable. As one of the most popular channels of communication, it's also a target for spammers and con artists. The way we interact socially has been revolutionized by the advent of these online communities. Making friends, staying in contact, and learning about people's lives has become simpler. While the use of social media has increased, so too have issues like phony accounts and online impersonation. This document provides a summary of the many approaches now in use to identify social media accounts that are not genuine. This article examines the available literature and draws comparisons between various studies. Intruders often create fake profiles on social networking sites in order to damage users' identities, steal their information, or invade their privacy. Therefore, one of the major challenges is figuring out if a given account is genuine or not.
Keywords: social media, Fake accounts, Machine learning algorithms, Comprehensive Review.

I. INTRODUCTION

The growth of social media in recent years has been phenomenal. It's crucial for marketing firms that want to build their brand by attracting new customers and supporters. Facebook and other forms of social media [2] have grown more ubiquitous and important in the modern world. Social media is utilized for a variety of purposes beyond simple communication, including advertising and brand promotion. At first glance, an account's popularity may be gauged by looking at metrics like follower count or the amount of likes, comments, or views on posted material. While social media [4] provide many benefits to our daily lives, they also raise some serious concerns that must be resolved. Fake accounts that have the appearance of being generated on behalf of organizations or individuals can damage reputation and reduce the number of likes and followers of individuals by exploiting legitimate concerns about privacy, online abuse, misuse, bullying, etc. However, it is predicted that the most harm would be done by the establishment of bogus accounts.

Fake social media profiles may be created for a variety of purposes, some of which are mentioned in [12].

Here are a few examples of why some individuals create phony accounts:

Social engineering (A) Online impersonation (B) Campaigning and Promotion

An Invasion of Privacy, Etc.

In most cases, spammers on social media are really legitimate people. Because of this, it is difficult to distinguish them from legitimate users. Furthermore, low-cost automated techniques are still accessible to fraudsters. gaining trust and credibility while making it difficult for the great majority of social media users to understand. Identifying false accounts on social media is a classification issue in which legitimate users are reliably distinguished from imposters by use of shared characteristics. A person's identity is something about them that exists independently of them.

A person's name is a typical instance of this. One more instance is a passport. Information such as a person's name, date of birth, place of birth, citizenship, digital fingerprints, and digital image are included. Each piece of [7] identification should only ever be used to relate to a single person at most. It is still possible for the same individual to use different names and identification numbers for different purposes. Authorities inside a nation-state will verify your identity if you claim to be someone else. One such example is a contemporary passport. Authorities certify that the photograph, fingerprints, name, birth date, etc. all belong to the same individual. A person's profile on a social networking site is usually the only way to tell who they are. This typically consists of a picture, a name, an address, and maybe a birth date. However, the platforms do not verify the

 ISSN: 2582-8150

profile's claimed subject's true identity or exert any control over its content. The alternative is that it is a person impersonating another. It's a case of adopting a fictitious persona.

Several methods of detection have been presented [18], broadly divided into three classes:

There are three stages: 1) learning characteristics, 2) social network-based, and 3) optimization.

When it comes to detecting false profiles, the performance of the first two types of algorithms that just involve features learning or social network knowledge falls short. The third strategy takes into account both feature data and social network information to discover a superior model.

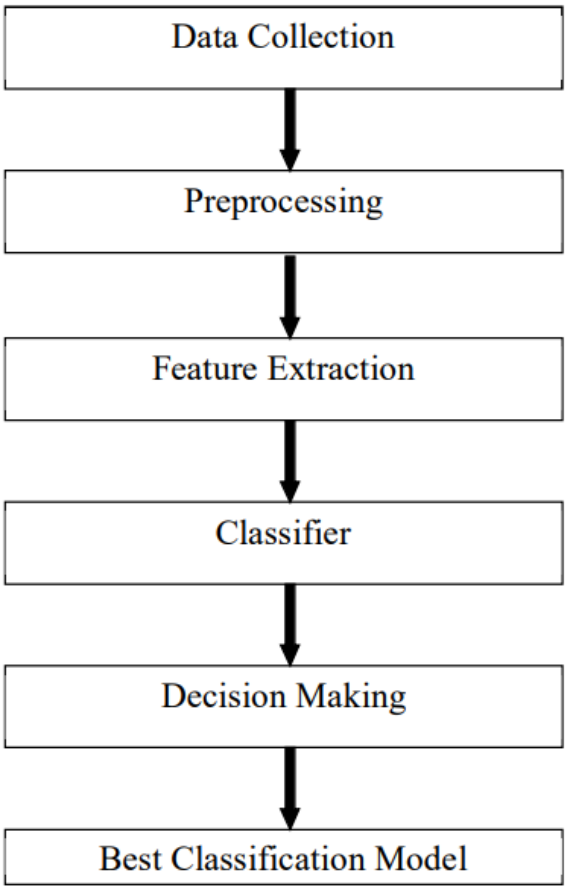

Figure 1: Research Steps involved in Fake Account Detection in social media.

Advanced Persistent Threats (APTs), or well-planned, long-term, and subtle efforts to compromise governmental, non-governmental, and corporate targets, rely heavily on false identities [7]. Spamming, exaggerating the number of users on a promotional app, etc. are just a few examples of how false identities may do damage.

Spammers' ability to assume false personae for various online activities is a major issue [9]. The spreading of rumors with the intent of changing the course of an institution or perhaps a whole society is one such goal.

There are telltale signs of fake accounts that may be easily recognized. [2] Describe the following as being indicative of a phony account:

a) Have a sizable fan base but a modest number of followers.

Preprocessing Collected Data

Using a Classifier to Extract Features

b) No profile pictures

c) weird user names

d) Excessive "liking"

II. PREVIOUS WORK

Recent studies have added a great deal of effort to the study of a distinct facet of social media by focusing on relevant research concerns.

Researchers Yasyn Elyusufi and colleagues. [2] This research offered a method for identifying a fake profile on a social networking site with very little information about the profile itself. Independently utilizing a supervised learning method on datasets that included both false and real users, the suggested model was trained.

The accuracy of predictions was improved by using an ensemble classifier. In this study, we use three different supervised machine learning techniques. It is possible to distinguish between fake and real profiles with the help of Random Forest, Decision Tree, and Naive Bayes. The results reveal that the Random Forest algorithm outperforms the alternative Algorithm with a 99.64% accuracy rate [4].

Dr. Farhan Nurdiatama Pakaya and coworkers. [5] In this study, we created a classification algorithm that uses just tweets from an account to determine if it is real or phony. Logistic Regression, ADA Boost, XG Boost, and Random Forest are the four methods employed in this study.

This study's examination of models suggests that XG Boost with tf- idf features is the best model for binary classification, whereas world2vec features is the best model for multiclass classification. A high of 95.5% accuracy was attained in this effort. According to [1]

 ISSN: 2582-8150

Fatih Cagatay Akyon and M. Esat Kalfaoglu [3]. In this study, we examine the challenge of spotting fraudulent and automated accounts that contribute to a sham cooperation on Instagram as one of two possible classes. This research offered a derived function for identifying automated and phony accounts, as well as a genetic algorithm-based cost-sensitive feature extraction approach for selecting appropriate characteristics for automated account classification.

Several machines Learning methods, including Naive Bayes, logistic regression, support Vector machines, and neural networks, are used to identify false and automated accounts. The highest F1 score was achieved by SVMs and neural networks. The best F1 score was achieved by the neural network (95%), whereas SVM only managed 86%. Mohammad B. A. Albayati and A. A. Altamimi [2019] A methodology for identifying fake profiles was proposed, and it makes use of many different data mining strategies. Twelve behavioral and non-behavioral discriminating profile traits were used in conjunction with supervised (ID3 decision tree, K-NN, and SVM) and unsupervised (K-Means, K-Medoids) machine learning methods. ID3's detection accuracy was determined to be 97.76% based on the findings. To combat the prevalence of phony social media profiles, S.P. Maniraj et al.[2019] developed a novel approach. The gradient boosting approach we're discussing here makes use of a decision tree with three categories.

Comment spam, fake activity, and low engagement rates are some examples of these aspects. In this work, we focus on using interaction rate and artificial activity as primary indicators of a phony account. Reference: [11] Zulfikar Alom et al. suggested a more robust set of tools for spotting Twitter trolls. It used seven distinct machine learning techniques, including K-NN, Decision Tree, Naive Bayes, Random Forest, Logistic Regression, Support Vector, and Extreme Gradient Boosting (XG-Boost), to display graph-based and tweet-based attributes. According to the results, the Random Forest algorithm outperforms the competition with an accuracy of 91%. Reference: [6] Mohd Fazil, Muhammad Abulaish [2018] In this research, we explored an approach for the automatic classification of spammers that integrates previously established community-based criteria with other categories of information including metadata, content-based interaction, and content-based qualities.

Because disregarding the contents of followers and the features of metadata was impractical, people have been categorized based on their interactions with their followers. In this experiment, we used a real-world dataset that included both valid users and spammers, who could be separated by 19 preexisting characteristics, 6 freshly specified features, and 2 redefined features.

Metadata categorization was the least efficient method, but interaction-based and community-based classifications were both shown to be useful. [3]

El Azab Ahmed, et al.[2016] In this study, we provide a categorization scheme for spotting Twitter bots. The research narrows down the elements that matter most in determining whether or not a Twitter account is phony, and then uses a variety of categorization methods to choose which ones to use. In this study, we employ many different categorization strategies, including Random Forest, Decision Tree, Naive Bayes, neural network, and Support vector machine. [10]

III. COMPARATIVE STUDY

In this study, comparison of different existing techniques and accuracy score is shown below :

Table 1.Comparison table of existing techniques and accuracy

Ref.	Year	Author	Technique	Accuracy
[19]	2015	Cao Xiao et al	Random forest, Support Vector Machine, and Logistic regression	95%
[8]	2017	Ashraf Khalil et al	Support Vector Machine, Simple Logistic , Instance-Based classifier using 1 nearest neighbor	98.74 %
[9]	2017	Buket Erşahin et al	Naïve Bayes	90.41 %
[6]	2018	Zulfikar Alom et al	Decision Tree, Naive Bayes , Random Forest, Logistic Regression,K-NN, Support Vector Machine and Extreme Gradient Boosting (XG-Boost).	91%
[21]	2018	Sarah Khaled et al	Support Vector Machine, Neural Network, SVM-NN	98%
[1]	2019	Farhan Nurdiatama Pakaya et al	Logistic Regression, ADA Boost, XG Boost, Random Forest	95.5%
[16]	2019	Hakimi A.N. et al	K-NN, SVM, NN	82%
[4]	2020	Yasyn Elyusufi et al	Random Forest, Decision Tree and Naive Bayes	99.64%

IV. DESCRIPTION OF THE DATASET

Following is a list of the chosen basic characteristics included in the dataset we want to employ in our study.

A. Whether or not a profile picture exists;

B. The profile's username;
C. Whether or not the profile is private; and
D. The account's total number of followers.
E. The number of people that are following the account.
F. Whether or not the account has a web address.
G. Publication in the Media

V. CONCEPTUAL PLAN

Our goal is to accurately detect the fraudulent account on social media using the fewest available characteristics. First, we employ a set of classification algorithms based on our findings on the most important criteria affecting the accurate detection of bogus accounts. The primary goal of this study is to develop an algorithm that outperforms the current one and enhances the effectiveness of the existing system.

VI. EXPECTED OUTCOME

This article provides a summary of previous studies that have investigated the use of classification methods for the identification of bogus accounts on social media. Support Vector Machine, Random Forest, Logistic Regression, Decision Tree, Naive Bayes, Neural Network, XG-Boost, ADA Boost, and K-NN are only few of the categorization algorithms utilized in various studies. Based on our research, Random Forest performs the best in terms of accuracy (99.64%) and is one of the most popular machine learning algorithms because of its versatility (it can be used for both regression and classification issues).

REFERENCES

[1] F. N. Pakaya, M. O. Ibrohim and I. Budi, "Malicious Account Detection on Twitter Based on Tweet Account Features using Machine Learning," 2019 Fourth International Conference on Informatics and Computing (ICIC), Semarang, Indonesia, 2019, pp. 1-5, doi: 10.1109/ICIC47613.2019.8985840.

[2] F. C. Akyon and M. Esat Kalfaoglu, "Instagram Fake and Automated Account Detection," 2019 Innovations in Intelligent Systems and Applications Conference (ASYU), Izmir, Turkey, 2019, pp. 1-7, doi: 10.1109/ASYU48272.2019.8946437.

[3] M. Fazil and M. Abulaish, "A Hybrid Approach for Detecting Automated Spammers in Twitter," in IEEE Transactions on Information Forensics and Security, vol. 13, no. 11, pp. 2707-2719, Nov. 2018, doi: 10.1109/TIFS.2018.2825958.

[4] Elyusufi, Yasyn & Elyusufi, Zakaria & M'hamed, Aït Kbir. (2020). Social Networks Fake Profiles Detection Using Machine Learning Algorithms. 10.1007/978-3-030-37629-1_3.

[5] Albayati, Mohammed & Altamimi, Ahmad. (2019). Identifying Fake Facebook Profiles Using Data Mining Techniques. Journal of ICT Research and Applications. 13. 107-117. 10.5614/itbj.ict.res.appl.2019.13.2.2.

[6] Z. Alom, B. Carminati and E. Ferrari, "Detecting Spam Accounts on Twitter," 2018 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM), Barcelona, 2018, pp. 1191-1198, doi: 10.1109/ASONAM.2018.8508495.

[7] Romanov, Aleksei & Semenov, Alexander & Mazhelis, Oleksiy & Veijalainen, Jari. (2017). Detection of Fake Profiles in Social Media - Literature Review. 363-369. 10.5220/0006362103630369.

[8] Ashraf Khalil, Hassan Hajjdiab, and Nabeel Al-Qirim, "Detecting Fake Followers in Twitter: A Machine Learning Approach," International Journal of Machine Learning and Computing vol. 7, no. 6, pp. 198-202, 2017.

[9] B. Erşahin, Ö. Aktaş, D. Kılınç and C. Akyol, "Twitter fake account detection," 2017 International Conference on Computer Science and Engineering (UBMK), Antalya, 2017, pp. 388-392, doi: 10.1109/UBMK.2017.8093420.

[10] Elazab, Ahmed & Mahmood, Mahmood & Hefny, Hesham. (2016). Fake Account Detection in Twitter Based on Minimum Weighted Feature set. International Journal of Computer, Electrical,

Automation, Control and Information Engineering Vol:10, No:1, 2016.

[11] S. P. Maniraj, Harie Krishnan G, Surya T, Pranav R (2019). Fake Account Detection using Machine Learning and Data Science. International Journal of Innovative Technology and Exploring Engineering (IJITEE) ISSN: 2278-3075, Volume-9 Issue-1, November, 2019.

[12] Bharat Sampatrao Borkar, Dr. Rajesh Purohit (2019).Recognisation of fake profiles in social media. Department of Computer Science & Engineering School of Engineering & Technology, Suresh Gyan Vihar University, Jagatpura, Volume-9 Issue-2,2019.

[13] Adikari, S. and K. Dutta. "Identifying Fake Profiles in LinkedIn." ArXiv abs/2006.01381 (2014).

[14] Devakunchari Ramalingam, Valliyammai Chinnaiah,"Fake profile detection techniques in large-scale online social networks: A comprehensive review" ,Computers & Electrical Engineering, Volume 65,2018.

[15] K. Zarei, R. Farahbakhsh and N. Crespi, "Typification of Impersonated Accounts on Instagram," 2019 IEEE 38th International Performance Computing and Communications Conference (IPCCC), London, United Kingdom, 2019, pp. 1-6, doi: 10.1109/IPCCC47392.2019.8958763.

[16] Hakimi, A.N., Ramli, S., Wook, M., Zainudin, N.M., Hasbullah, N.A., Wahab, N., & Afiza, M.R. (2019). Identifying Fake Account in Facebook Using Machine Learning. IVIC.

[17] F. Masood et al., "Spammer Detection and Fake User Identification on Social Networks," in IEEE Access, vol. 7, pp. 68140-68152, 2019, doi: 10.1109/ACCESS.2019.2918196.

[18] H. Shen and X. Liu, "Detecting Spammers on Twitter Based on Content and Social Interaction," 2015 International Conference on Network and Information Systems for Computers, Wuhan, 2015, pp. 413-417, doi: 10.1109/ICNISC.2015.82.

[19] Xiao, Cao et al. "Detecting Clusters of Fake Accounts in Online Social Networks." AISec '15 (2015).

[20] N. Singh, T. Sharma, A. Thakral and T. Choudhury, "Detection of Fake Profile in Online Social Networks Using Machine Learning," 2018 International Conference on Advances in Computing and Communication Engineering (ICACCE), Paris, 2018, pp. 231-234, doi: 10.1109/ICACCE.2018.8441713.

[21] S. Khaled, N. El-Tazi and H. M. O. Mokhtar, "Detecting Fake Accounts on Social Media," 2018 IEEE International Conference on Big Data (Big Data), Seattle, WA, USA, 2018, pp. 3672-3681, doi: 10.1109/BigData.2018.8621913.