International Journal of Electrical and Computer Engineering (IJECE)
Vol. 13, No. 3, June 2023, pp. 2962~2971
ISSN: 2088-8708, DOI: 10.11591/ijece.v13i3.pp2962-2971

Exploring machine learning techniques for fake profile detection in online social networks

Bharti, Nasib Singh Gill, Preeti Gulia
Department of Computer Science and Applications, Maharshi Dayanand University, Rohtak, India

Article Info

Article history:

Received May 27, 2022
Revised Jul 20, 2022
Accepted Aug 18, 2022

Keywords:

Fake profile
Machine learning
Online social network
Security threats
Unsupervised learning

ABSTRACT

The online social network is the largest network, more than 4 billion users use social media and with its rapid growth, the risk of maintaining the integrity of data has tremendously increased. There are several kinds of security challenges in online social networks (OSNs). Many abominable behaviors try to hack social sites and misuse the data available on these sites. Therefore, protection against such behaviors has become an essential requirement. Though there are many types of security threats in online social networks but, one of the significant threats is the fake profile. Fake profiles are created intentionally with certain motives, and such profiles may be targeted to steal or acquire sensitive information and/or spread rumors on online social networks with specific motives. Fake profiles are primarily used to steal or extract information by means of friendly interaction online and/or misusing online data available on social sites. Thus, fake profile detection in social media networks is attracting the attention of researchers. This paper aims to discuss various machine learning (ML) methods used by researchers for fake profile detection to explore the further possibility of improvising the machine learning models for speedy results.

Corresponding Author:

Bharti
Department of Computer Science and Applications, Maharshi Dayanand University
Rohtak, Haryana, 124001, India
Email: bharti.rs.dcsa@mdurohtak.ac.in

1. INTRODUCTION

Online social network is the most heard term used these days at every place. With the growth in technology, especially the internet, the craze for online social networks (OSNs) is increasing day by day. OSN are transforming how individuals communicate with one another [1]. About 4 billion people use different social media sites to connect with friends, family, and professional colleagues. So, risks of maintaining the privacy and security of users arise when the user's uploaded content are multimedia such as photos, and videos, and this information can be viral for a specific purpose [2]. With the rapid growth in technology, there is a growth in the number of users who use social media platforms. Billions of users have accounts on these sites. Some users create accounts on these sites and, for unethical purposes, hide their identities. Such user accounts are called fake profiles. Some people create fake accounts only for using social media for personal use like entertainment, education, and news. Still, there are some other users who hide their identity with mischievous aims. Such accounts are hazardous to our society. Detecting such profiles is essential in terms of security. Only a few researches have been done to identify fake profiles on social media platforms. Various machine learning (ML) methods are used to do this task [3].

The paper is further organized into the following sections: section 2 gives the idea about online social network where a brief discussion is made about social media. Section 3 represents online social network

Security Threats in which various security issues are discussed. Section 4 deals with the concept of ML. Section 5 specifies the role of ML in fake profile detection. Section 6 presents different challenges faced during threat detection and section 7 is about the conclusion and future scope of the study.

2. ONLINE SOCIAL NETWORK

Online social networks (OSN) are used significantly in the current scenario with the availability of the internet as technology is improved worldwide. The internet and web 3.0 types of machinery have made it easier to access online social networking sites like Facebook, Twitter, Instagram, and LinkedIn. As a result, people share their opinions and feelings about a wide range of topics with each other on these social media sites [4]. Figure 1 depicts the number of users of the most prominent social networking sites in January 2022. The figure data show how broad the scope of OSN is. With 2.91 billion monthly active users, Facebook, the global leader in this industry, is in the lead. YouTube is the second leader, and WhatsApp is at 3rd leading Instagram.

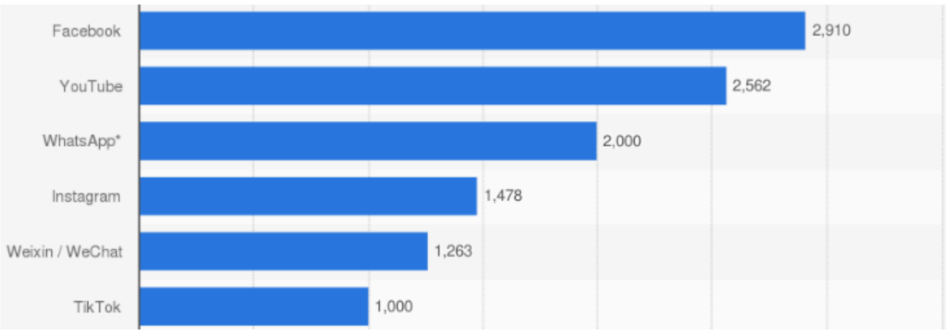

Figure 1. Number of active users in million on social media [5]

3. ONLINE SOCIAL NETWORK SECURITY THREATS

OSN are producing a big amount of data every single day. With the production of data, the risk of maintaining the security of this data is also increased. Many attackers are attracted to attack this data. So, there is always a need to protect the data on these sites. Social media security threats are the risks that occur due to sharing of data on online social media platforms and misuse by unauthorized users for malicious purposes. Such threats are a risk to the government as well as normal people. Many types of attacks are tried in the last two decades. Each and every day, a new attack is found on online social media. So, detecting such types of attacks is the biggest demand of today’s society. Figure 2 depicts the security threats in OSN.

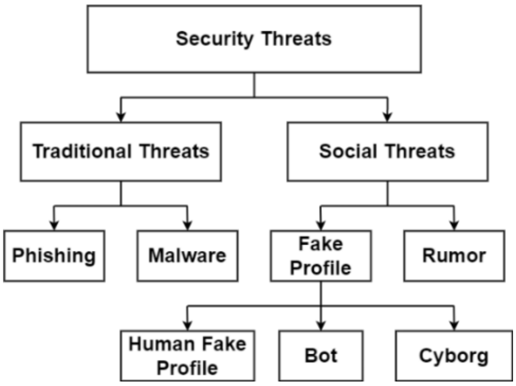

Figure 2. Types of security threats on social media

3.1. Traditional threats

In traditional social media security threats, traditional methods of security threats are used to attack users on online social media platforms. Some examples of traditional threats are phishing, and malware [2]. In phishing, money is stolen from targeted users. In phishing, the method used by fish catchers is used to trap the targeted users using attractive offers. Malware is the short form of ‘Malicious Software’ in which a computer program is used to attack and slow down the processing of the targeted computer system.

3.1.1. Phishing

Phishing is an attack where attackers attack using fake websites and emails. They create a fake website that looks the same as the original one. Sometimes attackers use social media sites to attack the users. In this method, they collect user information and then send a fake message which appears original. In this message, they ask users about their bank details. Users who are not well aware of this, log in on fake web links sent by attackers and after this, they are trapped by attackers. To perform such type of attacks, attackers use different techniques, such as an attractive advertisement like “Click here to see a famous actor naked”, and after clicking on the web link user become the victim of a phishing attack [2]. Now phishers have begun to use social media sites such as Twitter, and Facebook to disseminate phishing scams. Twitter is a popular microblogging platform where users can send 140-character messages called tweets. It has around 10 crore daily active users who send out approximately 20 crore tweets. Because of a large amount of information available, phishers have begun to utilize Twitter to spread phishing. Furthermore, unlike emails, phishing on Twitter is difficult to detect due to the rapid dissemination of phishing links in the network, the small size of the content, and the use of URLs [6].

3.1.2. Malware

Malware is spreading today due to its widespread use in the OSN, and it causes a variety of problems. One of the most common social network analysis (SNA) issues is detecting this malware. Social network services, in general, are made up of links between different user systems. As a result, malware can readily spread across users' computers via these links [2]. Malware is software that is meant to cause harm to a computer, server, client, or network, leak private information, obtain unauthorized access to information or systems, refuse users access to information, or mistakenly jeopardize a user's computer security and privacy. Malware is a severe threat to consumers and businesses. Viruses, worms, Trojan horses, and bots, among other things, are referred to as malware. There are many different types of malwares, each with its own method of infecting and spreading across computers. Malware can infect computers by being packaged with other programs or by being embedded as a macro in files. It can also infect a system by exploiting a known security flaw, such as a hole in an operating system, a network device, or a browser [7]. In 2005, the MySpace Samy worm was one of the first social media attacks. This is the first time an active worm has been found in OSNs. After 20 hours, it had infected over a million people's systems. MySpace had to close the site two days later to rectify the problem. Samy exploited a security flaw in the MySpace Web application program's cross-site script [8].

3.2. Social threats

In social threats, a social relationship is created with the user to be attacked. After that, a network is formed with some illegal motives to spread criminal activities such as pornography, cyber harassment, and spying. Some examples of social threats are fake profiles, and rumors [2]. In rumors, false information is shared on social media sites to attract the audience toward some specific topics. This method is mostly used by politicians to defame the opposite party candidates.

3.2.1. Fake profile

Users who hide their identity show themselves as some other users are called fake users, and the accounts held by such users on social media are called fake profiles. Not every fake account is a blemish on society. Some users create fake accounts for certain reasons and use them only for a limited period of time to fulfill their aims. But there are a number of accounts called fake accounts that have harmful intentions for society. Identifying such accounts is requisite for our community. Figure 3 depicts about the types of fake profiles in online social media.

Fake human profiles are the social media accounts created by humans for malicious purposes. They spread fake news and are handled by humans on online social media. These accounts are generated as well as used by human beings. Bots are computer-generated fake user profiles on online social media. The word 'Bot' is the short form of 'Robot'. It is a program that does the tasks repeatedly to perform a specific purpose. It mimics human behavior. Bots are used on social media to steal the personal information of celebrities, and defame a particular user. These are created by computers as well as operated by the computer to spread fake information. Some examples of bots are chatbots, Google bots, Social bots, and Malicious bots [9], [10]. Cyborg is a mixture of human fake profiles and bots. These profiles are created by humans but, after that, are used by bots [3]. These accounts are mostly used by politicians to libel opposite-party candidates. Sometimes such accounts are used to generate fake news on social media to attract the audience.

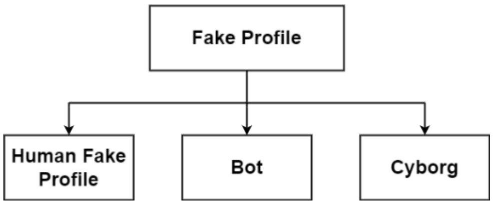

Figure 3. Types of fake profile

3.2.2. Rumor

Rumor is an OSN challenge that involves the transmission of false information, i.e., information manipulation. There are numerous definitions for the problem. This issue is defined by the Oxford Dictionary as a currently circulating rumor or report of dubious or disputed truth. A rumor is classified as an unconfirmed information piece. Therefore, it is impossible to know whether it is real or not while it is circulating. Rumor is described as unverified when there is no supporting proof, no official confirmation from approved sources, or when there are no reputable sources in a given situation [11]. Cyber-attacks such as spamming can be detected using rumors. Some rumor detection algorithms have analyzed and identified spammers using tweet posters. Spammers are more likely to spread false information and gain followers who can deceive them if they manage their tweeting activities [12].

4. MACHINE LEARNING

It is a branch of artificial intelligence (AI). In it, a machine is made to learn by itself without using any external programming or user interrupt. Machine learning plays a significant role in every field of research in computer science. There is no single field where machine learning is not used. Machine learning is not just a term, but it is making the machine learn how to solve a particular problem with or without human interaction. Machine learning plays a significant role in network security. We can use methods like support vector machine (SVM), random forest, and naive Bayes, to solve many security issues, especially in fake profile detection in online social media platforms. Figure 4 tells about the type of machine learning used in social media.

Supervised learning predicts future events based on past experience. In supervised learning, the predicted output is known in advance. It ensures that the model avoids overfitting and underfitting. unsupervised learning is a machine learning method for finding patterns in data. In an unsupervised algorithm, data is not labeled, which means only the input is given, and there is no corresponding output. In unsupervised learning, the algorithm discovers the pattern in data. Here the output is not predicted in advance. Reinforcement Learning is behavioral machine learning which uses the hit and trial method to produce output. In Reinforcement Learning, the predicted output is partially known. Reinforcement learning mimics supervised learning. Unlike supervised learning, there are no training labeled datasets. In reinforcement learning, agents learn from their past experiences. Since many machine learning methods are used in fake profile detection. Some of the important machine learning methods used in fake profile detection can be seen in subsections 4.1 to 4.3.

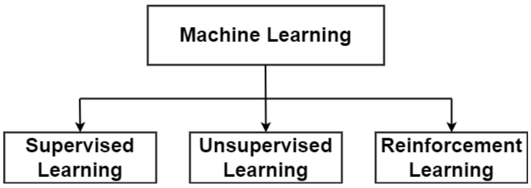

Figure 4. Types of machine learning [13]

4.1. K-nearest neighbor

It is a supervised learning method used for classification issues. It is widely used in pattern recognition, data mining, and intrusion detection [14]. It is also used for regression. It is also called a lazy learner algorithm means it does not learn from past experience. It does not make any assumptions about data, so it is called a non-parametric algorithm [15]. It classifies the data based on the parameter of similarity. In Figure 5 new data item is to be put in category 1 using k-nearest neighbor (KNN) based on similarity measures.

4.2. Random forest

It is a supervised learning strategy for classification and regression problems. For different samples, it builds decision trees. It handles continuous variables in a regression and categorical variables in classification [16]. The random forest contains many decision trees. More trees in the forest mean more precise output. It decreases overfitting in the data set. In this algorithm, every feature is selected at random, so it is called a random forest [17]. Figure 6 represents the flowchart to buy a new phone using random forest. In Figure 6 a random forest is created to buy a new smartphone based on some conditions like price, and 5G technology. After checking all conditions final decision is taken on whether to buy a new smartphone or not.

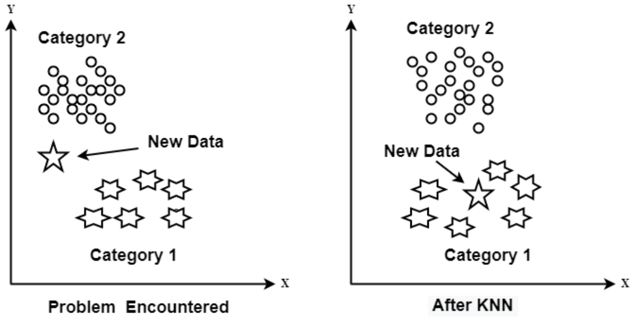

Figure 5. K nearest neighbor example [15]

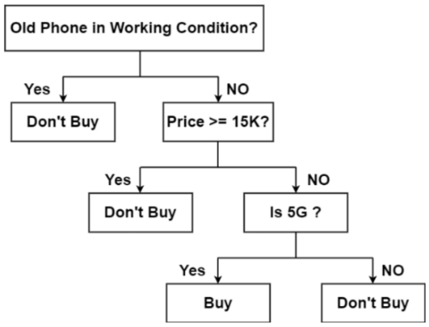

Figure 6. Random forest example

4.3. Naive Bayes

It is a supervised learning method that is based on the 'Bayes' theorem. It is used for classification problems. It is mostly used in text classification. It is a probabilistic classifier that is used to predict based on probability. Some examples are spam filtration, and sentiment analysis. The word ‘naive Bayes’ is made of 2 words, 'naive' and 'Bayes’ where naive represents that one feature is not dependent on other features. Bayes represents ‘Bayes’ theorem [18].

Bayes theorem: It determines the probability of the hypothesis.
The formula is as: $P(A|B) = [\ P(B|A)P(A)\ ] / P(B)$
where A = Hypothesis, B = Observed Event.
$P(A|B)$ = Probability of A on B.
$P(B|A)$ = Probability of evidence when $P(A)$ = True.
$P(A)$ = Probability of hypothesis A.
$P(B)$ = Probability of B.

5. ROLE OF MACHINE LEARNING IN FAKE PROFILE DETECTION

Since a lot of research work has been done in fake profile detection, many machine learning methods have been used by different researchers to identify such identities. Figure 7 shows the work flow of machine learning technique in fake profile detection. Figure 7 represents the step-by-step procedure to detect the fake profiles using machine learning. First of all, raw data is collected then features are extracted based on some conditions. Then machine learning based classifiers will decide whether the selected accounts are fake or real.

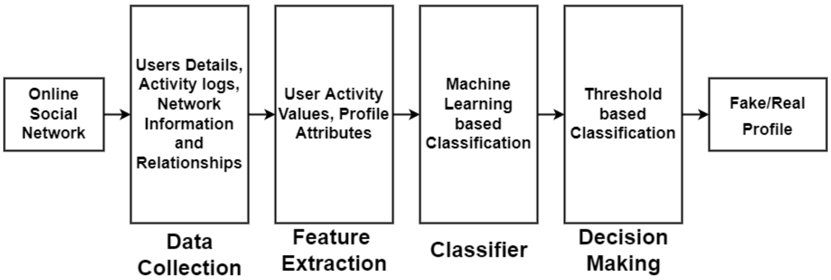
Online Social Network
Users Details, Activity logs, Network Information and Relationships
User Activity Values, Profile Attributes
Machine Learning based Classification
Threshold based Classification
Fake/Real Profile
Data Collection
Feature Extraction
Classifier
Decision Making

Figure 7. Machine learning technique work flow [19]

Some of the key findings in the related research areas are as follows: Kolomeets *et al.* [20] focused on the case for 'fake profile detection' while profile was locked due to secrecy situations and had to find fake profiles based on the friend list. Machine learning and statistical methods were used. For data source VKontakte (VK) was used. Statistics, Benford's law, and the Gini index were used for feature construction. The experiment was performed in two series. In the first series, they train classifiers, and in the second series, they compared the features. The RF classifier gave the maximum result, with receiver operator characteristic (ROC) AUC >0.9 and FPR<0.3. Lingam *et al.* [21] considered users' temporal behavioral features "average number of tweets posted per day, longest user session time without a break, and percentage of dropped followers per week" on Twitter. "Social honeypot" and "the fake project" datasets were used. Finally, they concluded 'learning with example patterns using supervised learning' does not provide correct results where bot behavior changes dynamically. Muñoz and Paul [22], proposed some machine learning methods for 'fake profile detection', especially unsupervised learning on Instagram. Seventeen metadata features from real and fake accounts were used in the dataset. The random forest method provided 96% accuracy. However, the dataset can be increased to provide more satisfactory results.

Lê *et al.* [23] proposed an "empirical ranking scheme" involving "graph-based and feature-based approaches" to detect fake accounts on Facebook. SVM and SybilWalk algorithm were used. Ten thousand Vietnamese Facebook accounts were used for the proposed work. The precision of fake accounts was 0.8, and real accounts was 0.92, so the approach was proved good after about 50 iterations. Punkamol and Marukatat [24] proposed a method that identified account cloning in Twitter based on user profiles, friends, "follower networks," and posting behaviors. "Twitter Crawler, Attribute Extractor, and Cloning Detector" were parts of the framework. Research results provided that completely cloning of user accounts was not possible, and identifying clone accounts based on writing style, behavior, and was easy. In this paper, the decision tree method gave the best result. Sowmya and Chatterjee [25], proposed a method that detected Clone accounts based on a set of rules on Twitter. Similarity Measures and C4.5 decision tree methods were used for cloning detection. The similarity of Attributes and Similarity of Network relationships were used. C4.5 builds a decision tree to identify clones. At last, a comparison was made between both methods. Similarity Measure provided better results than the C4.5 decision tree algorithm.

Roy and Chahar [26] did a survey to detect phony accounts on social media. The existing work was categorized into three groups: i) research using non-textual features, ii) research using textual features, and iii) research using both non-textual and textual features. They compare different existing methods of fake profile detection in this survey. They discussed many issues related to fake profile detection, but the main issue they found was the language dependence of programs used on online social media platforms. Now users on social media use mixed language, so it is difficult to understand the meaning of the message. Zarei *et al.* [27], suggested a model to isolate Impersonators on Instagram. The engagement of impersonators was studied in active and passive engagements. Politicians, News agencies, and Sports stars were targeted. Natural Language Processing was used to understand the comments. K-means, Gaussian Mixture Model, and Spectral Clustering algorithms were used, and engagement of Impersonators was successfully studied. Wani *et al.* [28], analyzed real and fake users on Facebook based on emotions. They trained their model based on 12 emotions using different machine learning approaches containing "SVM, Naive Bayes, JRip, and Random Forest". They used the 'Honeypots' technique to collect data from Facebook. They used the Python library for database construction and finally used scatter and boxplot graphs for feature analysis. They found that real profile posts contain more emotions than fake profiles.

Tiwari [29], reviewed many methods to identify fake accounts on social networks. He also discusses the various social network security-related issues and different techniques used by different authors to avoid these security threats. He discussed many Machine learning methods already used by different researchers to detect fake profiles and concluded 'social engineering' is the main threat in OSN. Gupta and Kaushal [30], fake profiles on Facebook (FB). Social activity-based learning methods were used to identify fake profiles on Facebook. naive Bayes, J48, random forest, random tree, REPTree, One R, and JRip algorithms were used. Finally, they concluded that user activities such as likes, comments, and shares, paid maximum to identify the fake accounts. Egele *et al.* [31], proposed a model to detect the composition of special high-profile accounts. COMPA detection system was designed which check message similarity on Facebook and Twitter. A statical model was used and behavioral features were studied to detect fake user profiles.

5.1. Comparative study and techniques analysis

Since a lot of research has been carried out related to the detection of fake accounts still. However, based on the literature reviewed and their comparative analysis several advantages and disadvantages have been explored, which inspires further work in the related area. These are presented in Table 1.

Table 1. Comparative analysis

Key Reference	Techniques Used	Advantages	Disadvantages
Lyu *et al.* [32]	Random forest SVM KNN	Sybil detection precision is excellent.	For more advancement, deep learning (DL) was not used.
Rezaimehr and Dadkhah [33]	C# Programming language	A New C# tool was designed which can help in dataset creation and detection of fake profiles.	The dataset created by the tool contains fake profiles intentionally.
Praveena and Vivekanandan [34]	ML Deep learning	Machine and deep learning both are used and compared at the same time.	The open research challenges have been portrayed to develop identification of shilling attacks in collaborative filtering-based recommender systems.
Caporusso *et al.* [35]	Generative adversarial networks Eye-tracking technology	Research results are useful in the healthcare industry.	The results of the individuals indicate no link with any of the other factors like demographic information, and familiarity with computers.
Patel *et al*. [36]	Supervised ML Unsupervised ML	Supervised ML and unsupervised ML both were used.	Bots were unable to distinguish between real and false profiles created by humans.
Pizarro [37]	SVM	10-fold cross-validation was utilized for training.	Only one technique was used.
Zarei *et al.* [38]	Unsupervised clustering	A complete analysis is provided about inner hidden clusters.	Supervised learning was not used in the proposed method.
Shu *et al.* [39]	Convolutional neural network Random forest	Fake profiles were used to detect fake news.	Users who shared, liked, and retweeted the posts were not included.
Chen *et al.* [40]	ML	Both trustworthy and untrustworthy users were analyzed.	Temporal features were not analyzed.
Das *et al.* [41]	A* Search Algorithm	Heuristic cost function proved good.	Only twitter dataset was used.

5.2. Datasets and result accuracy analysis

Since many different types of datasets are used by many researchers daily new datasets are created and used by different dataset scientists. All datasets have their unique quality. Research results do not depend only on the datasets but also depends on the methodology used by researchers too. Some datasets provide good accuracy in terms of false-positive rates some others provide good accuracy in terms of other factors. So, it is not crucial to say which dataset provides the best result with which technology. Some examples of these datasets used with different technologies giving accurate results in terms of different factors are summarized in Table 2.

Table 2. Dataset and result discussion

Key Reference	Dataset used	Most Feasible Technique	Result
Schler *et al.* [42]	Facebook	Neural Network	60 to 70 % with 0.01 p-value.
Morales *et al.* [43]	Aalto-University Dataset	Deep learning techniques	52.6 % accuracy.
Zarei *et al.* [44]	Politicians, Sport Starts, Musicians	Blend of synthetic minority over-sampling technique (SMOTE) and Random under-sampling algorithm	Accuracy=0.86, Precision= 0.85, Recall= 0.86, and F1= 0.85.
Kumar *et al.* [45]	spam.csv dataset from Kaggle	Classic classifiers (NB classifier, SVM, Decision tree, KNN) Ensemble learning (RF, Bagging, Boosting and AdaBoost classifier)	No Numerical value to accuracy is discussed.
Ebrahimiam and Kashef [46]	MovieLens 100K, Netflix	Deep-learning, convolutional neural networks	Accuracy and F-measure of up to 99 percent were achieved.
Ganguli *et al.* [47]	Twitter, LinkedIn	SMO-PolyKernel with decision trees, random forest	23% accuracy on Twitter using SMO, and using random forest on LinkedIn AUC=0.978, Recall: 0.900 were achieved.
Hajdu *et al.* [48]	Facebook	Artificial neural network	Only technique is applied results are not discussed.
Suarez-Tangil *et al.* [49]	Datingnmore.com Scamdigger.com	Machine learning	97% accuracy is received.
Sensonetti *et al.* [50]	Twitter	Neural network	Accuracy = 90% in differentiating false news from true news, and reliability = 92% is obtained in offline analysis of profiles.

6. CHALLENGES

Dataset collection and filtering are a big challenge in online social media. Here, only Twitter, Facebook, Instagram, and Netflix datasets are reviewed. Further, other social media sites like LinkedIn, Indeed, Naukri.com, and WhatsApp. may also be challenging. In recent years, the use of social media has increased for dating and life partner searches. Though detection of a fake profile is an open challenge, however, these databases may certainly contribute significantly to the detection of fake profiles.

7. CONCLUSION AND FUTURE SCOPE

OSN are the open platform for attackers and with the availability of huge data on these platforms, attackers are obviously attracted to such platforms. Detection of fake profiles is the most challenging task in the present-day digital universe. This paper throws light on fake accounts, security threats in online social media, and different machine learning techniques used by different researchers in this domain. Many issues regarding fake account detection on social media have been discussed. Relevant work done in this area including the dataset used, techniques used, and results in terms of performance have been reviewed in this paper. Machine learning seems to play a vital role in the detection of fake accounts. Many machine learning approaches are used by many researchers. Some of the significant research gaps have been identified, discussed, and analyzed, which leads to further research direction in the related area of detection of fake human accounts and profiles. Detailed analysis of the research review reveals that more approaches may be explored to bridge the gaps to help provide better results for fake profile detection. The paper also concludes that supervised machine learning is used by most researchers and the detection of bots is easier than human fake accounts. Therefore, there is a great scope for the detection of fake human accounts and profiles.

ACKNOWLEDGMENTS

The first author acknowledges University Grant Commission (UGC) India for providing Junior Research Fellowship (JRF) to carry out this work.

REFERENCES

[1] M. Conti, R. Poovendran, and M. Secchiero, “FakeBook: Detecting fake profiles in on-line social networks,” in *2012 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining*, Aug. 2012, pp. 1071–1078, doi: 10.1109/ASONAM.2012.185.

[2] S. Rathore, P. K. Sharma, V. Loia, Y.-S. Jeong, and J. H. Park, “Social network security: Issues, challenges, threats, and solutions,” *Information Sciences*, vol. 421, pp. 43–69, Dec. 2017, doi: 10.1016/j.ins.2017.08.063.

[3] E. Van Der Walt and J. Eloff, “Using machine learning to detect fake identities: Bots vs humans,” *IEEE Access*, vol. 6, pp. 6540–6549, 2018, doi: 10.1109/ACCESS.2018.2796018.

[4] U. Can and B. Alatas, “A new direction in social network analysis: Online social network analysis problems and applications,” *Physica A: Statistical Mechanics and its Applications*, vol. 535, Dec. 2019, doi: 10.1016/j.physa.2019.122372.

[5] Statista, “Most used social media 2021,” *Statista*. https://www.statista.com/statistics/272014/global-social-networks-ranked-by-number-of-users/ (accessed May 26, 2022).

[6] A. Aggarwal, A. Rajadesingan, and P. Kumaraguru, “PhishAri: Automatic realtime phishing detection on twitter,” in *2012 eCrime Researchers Summit*, Oct. 2012, pp. 1–12, doi: 10.1109/eCrime.2012.6489521.

[7] Cisco, “What Is the difference: Viruses, worms, trojans, and bots?,” *Cisco Security*. https://tools.cisco.com/security/center/resources/virus_differences (accessed May 26, 2022).

[8] M. R. Faghani and H. Saidi, “Malware propagation in online social networks,” in *2009 4th International Conference on Malicious and Unwanted Software (MALWARE)*, Oct. 2009, pp. 8–14, doi: 10.1109/MALWARE.2009.5403023.

[9] Cloudflare, “What is a bot? | Bot definition,” *Cloudflare*. https://www.cloudflare.com/learning/bots/what-is-a-bot/ (accessed Jul. 06, 2022).

[10] Kaspersky, “What are bots? – Definition and explanation,” *www.kaspersky.com*, 2022. https://www.kaspersky.com/resource-center/definitions/what-are-bots (accessed Jul. 06, 2022).

[11] A. Zubiaga, A. Aker, K. Bontcheva, M. Liakata, and R. Procter, “Detection and resolution of rumours in social media,” *ACM Computing Surveys*, vol. 51, no. 2, pp. 1–36, Mar. 2019, doi: 10.1145/3161603.

[12] Y. K. Yang, K. Niu, and Z. He, “Exploiting the topology property of social network for rumor detection,” in *2015 12th International Joint Conference on Computer Science and Software Engineering (JCSSE)*, Jul. 2015, pp. 41–46, doi: 10.1109/JCSSE.2015.7219767.

[13] A. Sagu and N. S. Gill, “Machine learning techniques for securing IoT environment,” *International Journal of Innovative Technology and Exploring Engineering*, vol. 9, no. 4, pp. 977–982, Feb. 2020, doi: 10.35940/ijitee.D1209.029420.

[14] GeeksforGeeks, “K-Nearest Neighbours,” Geeksforgeeks. https://www.geeksforgeeks.org/k-nearest-neighbours/ (accessed Jul. 06, 2022).

[15] “K-nearest neighbor (KNN) algorithm for machine learning,” JavaTpoint. https://www.javatpoint.com/k-nearest-neighbor-algorithm-for-machine-learning. (Accessed Jul. 06, 2022).

[16] E R Sruthi, “Random forest | Introduction to random forest algorithm.” Analytics Vidhya, 2021. https://www.analyticsvidhya.com/blog/2021/06/understanding-random-forest/ (accessed Jul. 06, 2022).

[17] O. Mbaabu, “Introduction to random forest in machine learning,” *Engineering Education (EngEd) Program | Section*. 2020, https://www.section.io/engineering-education/introduction-to-random-forest-in-machine-learning (accessed Jul. 06, 2022).

[18] "Naive bayes classifier in machine learning," JavaTpoint. https://www.javatpoint.com/machine-learning-naive-bayes-classifier (accessed Jul. 06, 2022).

[19] D. Ramalingam and V. Chinnaiah, "Fake profile detection techniques in large-scale online social networks: A comprehensive review," *Computers & Electrical Engineering*, vol. 65, pp. 165–177, Jan. 2018, doi: 10.1016/j.compeleceng.2017.05.020.

[20] M. Kolomeets, O. Tushkanova, D. Levshun, and A. Chechulin, "Camouflaged bot detection using the friend list," in *2021 29th Euromicro International Conference on Parallel, Distributed and Network-Based Processing (PDP)*, Mar. 2021, pp. 253–259, doi: 10.1109/PDP52278.2021.00048.

[21] G. Lingam, R. R. Rout, D. V. L. N. Somayajulu, and S. K. Ghosh, "Particle swarm optimization on deep reinforcement learning for detecting social spam bots and spam-influential users in twitter network," *IEEE Systems Journal*, vol. 15, no. 2, pp. 2281–2292, Jun. 2021, doi: 10.1109/JSYST.2020.3034416.

[22] S. D. Munoz and E. P. G. Pinto, "A dataset for the detection of fake profiles on social networking services," in *2020 International Conference on Computational Science and Computational Intelligence (CSCI)*, Dec. 2020, pp. 230–237, doi: 10.1109/CSCI51800.2020.00046.

[23] N. C. Le, M.-T. Dao, H.-L. Nguyen, T.-N. Nguyen, and H. Vu, "An application of random walk on fake account detection problem: A hybrid approach," in *2020 RIVF International Conference on Computing and Communication Technologies (RIVF)*, Oct. 2020, pp. 1–6, doi: 10.1109/RIVF48685.2020.9140749.

[24] D. Punkamol and R. Marukatat, "Detection of account cloning in online social networks," in *2020 8th International Electrical Engineering Congress (iEECON)*, Mar. 2020, pp. 1–4, doi: 10.1109/iEECON48109.2020.229558.

[25] P. Sowmya and M. Chatterjee, "Detection of fake and clone accounts in Twitter using classification and distance measure algorithms," in *2020 International Conference on Communication and Signal Processing (ICCSP)*, Jul. 2020, pp. 67–70, doi: 10.1109/ICCSP48568.2020.9182353.

[26] P. K. Roy and S. Chahar, "Fake profile detection on social networking websites: A comprehensive review," *IEEE Transactions on Artificial Intelligence*, vol. 1, no. 3, pp. 271–285, Dec. 2020, doi: 10.1109/TAI.2021.3064901.

[27] K. Zarei, R. Farahbakhsh, and N. Crespi, "How impersonators exploit Instagram to generate fake engagement?," in *ICC 2020 - 2020 IEEE International Conference on Communications (ICC)*, Jun. 2020, pp. 1–6, doi: 10.1109/ICC40277.2020.9149431.

[28] M. A. Wani, N. Agarwal, S. Jabin, and S. Z. Hussain, "Analyzing real and fake users in Facebook network based on emotions," in *2019 11th International Conference on Communication Systems & Networks (COMSNETS)*, Jan. 2019, pp. 110–117, doi: 10.1109/COMSNETS.2019.8711124.

[29] V. Tiwari, "Analysis and detection of fake profile over social network," in *2017 International Conference on Computing, Communication and Automation (ICCCA)*, May 2017, pp. 175–179, doi: 10.1109/CCAA.2017.8229795.

[30] A. Gupta and R. Kaushal, "Towards detecting fake user accounts in facebook," in *2017 ISEA Asia Security and Privacy (ISEASP)*, Jan. 2017, pp. 1–6, doi: 10.1109/ISEASP.2017.7976996.

[31] M. Egele, G. Stringhini, C. Kruegel, and G. Vigna, "Towards detecting compromised accounts on social networks," *IEEE Transactions on Dependable and Secure Computing*, vol. 14, no. 4, pp. 447–460, Jul. 2017, doi: 10.1109/TDSC.2015.2479616.

[32] C. Lyu *et al.*, "Predictable model for detecting sybil attacks in mobile social networks," in *2021 IEEE Wireless Communications and Networking Conference (WCNC)*, Mar. 2021, pp. 1–6, doi: 10.1109/WCNC49053.2021.9417254.

[33] F. Rezaimehr and C. Dadkhah, "Injection shilling attack tool for recommender systems," in *2021 26th International Computer Conference, Computer Society of Iran (CSICC)*, Mar. 2021, pp. 1–4, doi: 10.1109/CSICC52343.2021.9420553.

[34] N. Praveena and K. Vivekanandan, "A study on shilling attack identification in SAN using collaborative filtering method based recommender systems," in *2021 International Conference on Computer Communication and Informatics (ICCCI)*, Jan. 2021, pp. 1–5, doi: 10.1109/ICCCI50826.2021.9402676.

[35] N. Caporusso, K. Zhang, and G. Carlson, "Using eye-tracking to study the authenticity of images produced by generative adversarial networks," in *2020 International Conference on Electrical, Communication, and Computer Engineering (ICECCE)*, Jun. 2020, pp. 1–6, doi: 10.1109/ICECCE49384.2020.9179472.

[36] K. Patel, S. Agrahari, and S. Srivastava, "Survey on fake profile detection on social sites by using machine learning algorithm," in *2020 8th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO)*, Jun. 2020, pp. 1236–1240, doi: 10.1109/ICRITO48877.2020.9197935.

[37] J. Pizarro, "Profiling bots and fake news spreaders at PAN'19 and PAN'20: Bots and gender profiling 2019, profiling fake news spreaders on Twitter 2020," in *2020 IEEE 7th International Conference on Data Science and Advanced Analytics (DSAA)*, Oct. 2020, pp. 626–630, doi: 10.1109/DSAA49011.2020.00088.

[38] K. Zarei, R. Farahbakhsh, and N. Crespi, "Typification of impersonated accounts on Instagram," in *2019 IEEE 38th International Performance Computing and Communications Conference (IPCCC)*, Oct. 2019, pp. 1–6, doi: 10.1109/IPCCC47392.2019.8958763.

[39] K. Shu, X. Zhou, S. Wang, R. Zafarani, and H. Liu, "The role of user profiles for fake news detection," in *Proceedings of the 2019 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining*, Aug. 2019, pp. 436–439, doi: 10.1145/3341161.3342927.

[40] X. Chen, Y. Yuan, L. Lu, and J. Yang, "A multidimensional trust evaluation framework for online social networks based on machine learning," *IEEE Access*, vol. 7, pp. 175499–175513, 2019, doi: 10.1109/ACCESS.2019.2957779.

[41] R. Das, G. Karmakar, and J. Kamruzzaman, "How much i can rely on you: measuring trustworthiness of a Twitter user," *IEEE Transactions on Dependable and Secure Computing*, vol. 18, no. 2, pp. 949–966, Mar. 2021, doi: 10.1109/TDSC.2019.2929782.

[42] J. Schler, E. Bonchek-Dokow, T. Vainstein, M. Gotam, and M. Teplitsky, "Profiling astroturfing Facebook users during three contiguous Israeli election periods," in *2020 IEEE International Conference on Big Data (Big Data)*, Dec. 2020, pp. 4331–4340, doi: 10.1109/BigData50022.2020.9378207.

[43] A. Morales *et al.*, "Keystroke biometrics in response to fake news propagation in a global pandemic," in *2020 IEEE 44th Annual Computers, Software, and Applications Conference (COMPSAC)*, Jul. 2020, pp. 1604–1609, doi: 10.1109/COMPSAC48688.2020.00-26.

[44] K. Zarei, R. Farahbakhsh, N. Crespi, and G. Tyson, "Impersonation on social media: A deep neural approach to identify ingenuine content," in *2020 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM)*, Dec. 2020, pp. 11–15, doi: 10.1109/ASONAM49781.2020.9381437.

[45] N. Kumar, S. Sonowal, and Nishant, "Email spam detection using machine learning algorithms," in *2020 Second International Conference on Inventive Research in Computing Applications (ICIRCA)*, Jul. 2020, pp. 108–113, doi: 10.1109/ICIRCA48905.2020.9183098.

[46] M. Ebrahimian and R. Kashef, "Efficient detection of shilling's attacks in collaborative filtering recommendation systems using deep learning nodels," in *2020 IEEE International Conference on Industrial Engineering and Engineering Management (IEEM)*, Dec. 2020, pp. 460–464, doi: 10.1109/IEEM45057.2020.9309965.

[47] R. Ganguli, A. Mehta, and S. Sen, "A survey on machine learning methodologies in social network analysis," in *2020 8th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO)*, Jun. 2020, pp. 484–489, doi: 10.1109/ICRITO48877.2020.9197984.

[48] G. Hajdu, Y. Minoso, R. Lopez, M. Acosta, and A. Elleithy, "Use of artificial neural networks to identify fake profiles," in *2019 IEEE Long Island Systems, Applications and Technology Conference (LISAT)*, May 2019, pp. 1–4, doi: 10.1109/LISAT.2019.8817330.

[49] G. Suarez-Tangil, M. Edwards, C. Peersman, G. Stringhini, A. Rashid, and M. Whitty, "Automatically dismantling online dating fraud," *IEEE Transactions on Information Forensics and Security*, vol. 15, pp. 1128–1137, 2020, doi: 10.1109/TIFS.2019.2930479.

[50] G. Sansonetti, F. Gasparetti, G. D'aniello, and A. Micarelli, "Unreliable users detection in social media: Deep learning techniques for automatic detection," *IEEE Access*, vol. 8, pp. 213154–213167, 2020, doi: 10.1109/ACCESS.2020.3040604.

BIOGRAPHIES OF AUTHORS

Bharti received the B.Sc. and M.Sc. degrees in computer science from Maharani Kishori Kanya Mahavidyalaya Hodal Palwal (Haryana) affiliated to Maharshi Dayanand University, Rohtak, Haryana, India, in 2015 and 2017, respectively. Currently, she is Ph.D. Research Scholar at the Department of Computer Science and Applications, Maharshi Dayanand University, Rohtak, Haryana, India. She is awarded Junior Research Fellowship in the National Eligibility Test (UGC-NET) in India for her research. She may be contacted at email: bharti.rs.dcsa@mdurohtak.ac.in.

Nasib Singh Gill holds Post-Doctoral research in Computer Science at Brunel University, West London during 2001-2002 and Ph.D. in Computer Science in 1996. He is a recipient of the Commonwealth Fellowship Award of the British Government for the Year 2001. Besides, he also has earned his MBA degree. He is currently Head, Department of Computer Science and Applications, M. D. University, Rohtak, India. He is also working as Director, Directorate of Distance Education as well as Director of Digital Learning Centre, M. D. University, Rohtak, Haryana. He is an active professional member of IETE, IAENG, and CSI. He has published more than 304 research papers and authored 5 popular books He has guided so far 12 Ph.D. scholars as well as guiding about 5 more scholars. His research interests primarily include – IoT, machine and deep learning, information and network security, data mining and data warehousing, NLP, and measurement of component-based systems. He can be contacted at email: nasib.gill@mdurohtak.ac.in.

Preeti Gulia received Ph.D. degree in computer science in 2013. She is currently working as Associate Professor at Department of Computer Science and Applications, M.D. University, Rohtak, India. She is serving the Department since 2009. She has published more than 65 research papers and articles in journals and conferences of National/ International repute including ACM, and Scopus. Her area of research includes data mining, big data, machine learning, deep learning, IoT, and software engineering. She is an active professional member of IAENG, CSI, and ACM. She is also serving as Editorial Board Member Active Reviewer of International/National Journals. She has guided one research scholar as well as guiding four Ph.D. research scholars from various research areas. She can be contacted at email: preeti@mdurohtak.ac.in.