International Journal of Innovative Research in Computer Science & Technology (IJIRCST)
ISSN (online): 2347-5552, Volume-11, Issue-2, March 2023
https://doi.org/10.55524/ijircst.2023.11.2.9
Article ID IRP1374, Pages 44-49
www.ijircst.org

A Review Article on Detection of Fake Profile on Social-Media

Shamim Ahmad[1], and Dr. Manish MadhavaTripathi[2]

[1]M.Tech. Scholar, Department of Computer Science and Engineering, Integral University, Lucknow, India
[2]Associate Professor, Department of Computer Science and Engineering, Integral University Lucknow, India

ABSTRACT- Nowadays, practically everyone, from a youngster to an adult, spends more time on online social media platforms, connecting with and exchanging information with individuals all over the world. The social network is becoming a popular means to communicate with people who live in different parts of the world. Because of the tremendous interconnectedness and information sharing enabled by the internet, social media platforms. This highlights the importance of establishing a system capable of detecting fake profiles on social media networks. There has been a lot of study done in this area utilising machine learning algorithms to identify fake profile, duplicate, spam, and bot accounts, and most of the fake profile accounts were effectively recognised using machine learning algorithms. This study discusses fake profile detection on social networks using machine learning.

KEYWORDS- Fake Profile, Fake Identities, Security Issues, Social Network Analysis, Machine Learning

I. INTRODUCTION

Every user of a social networking site has a profile and can communicate with friends, exchange updates, and network with new people. These social networks leverage Web 2.0 technology, which facilitates user communication. These social networking sites expand swiftly and alter how people connect with one another. Via the online community, people with similar interests can meet and form connections. social effect During your contemporary generation, online social connections have become integral to everyone's social life. Most people are connected to some of these websites. OSNs like Instagram, Facebook, Google+, Twitter, LinkedIn, and Pinterest, among others, are expanding as a result of their free memberships and cost-free information sharing with other users. and they have become an essential part of life in today's generation. There are several disadvantages to the expanding use of OSNs, including a higher likelihood of fraudulent profiles, identity theft, privacy breaches, etc. False profile development has increased in tandem with the growth in social network users. This has contributed to a rise in cybercrime over the past few years.

Hackers exploit these sites to disseminate rumours, hate speech, and false information and to earn money unlawfully by setting up several bogus profiles. Also, people are making fictitious profiles to further their own interests, such as obtaining referral bonuses or an increase in votes in online voting systems. Researchers use social bots, which are automated programmes, to carry out their tasks. The fake accounts exist anywhere on the internet, such as on social networks, online dating websites, discussion blogs, shopping websites, etc. As social media platforms are used more often, users and platform providers are becoming increasingly concerned about the issue of fraudulent profiles. To propagate false information, con individuals, or carry out other nefarious deeds, fake profiles might be made. By examining a variety of profile data and user behaviour, machine learning techniques can be utilised to identify these fraudulent profiles.

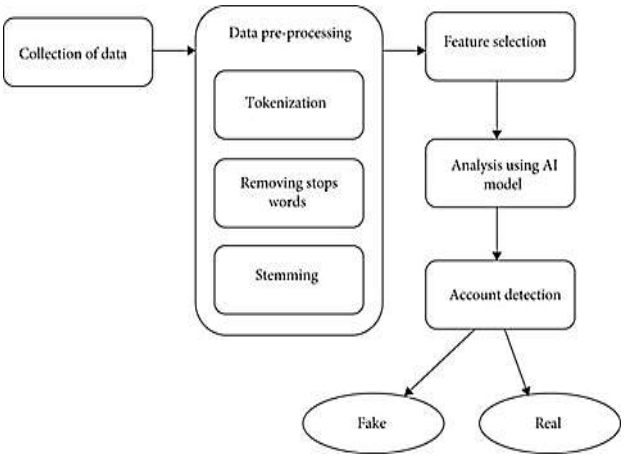

Figure 1: Analysing Model of Real Fake Detection

The detection of fake profiles using machine learning involves building a model that can classify a given profile as either genuine or fake. The model can be trained using a dataset of known fake and genuine profiles. The dataset can be collected using various methods, such as manual identification by experts, crawling social media platforms, or using crowd sourcing.

The features of the profiles that can be used for classification include the profile picture, name, location, age, gender, profile description, number of followers, friends, and posts, engagement rate, activity time, and content. Machine learning algorithms, such as decision trees, random forests, support vector machines, and neural networks, can be used to extract features from the profiles and classify them as genuine or fake.

Once the model is trained, it can be used to classify new profiles automatically. However, the performance of the model depends on the quality of the dataset, the choice of features, and the algorithm used. The model may also need to be updated periodically to adapt to new techniques used by fake profile creators.

Overall, the use of machine learning techniques can help in the early detection and prevention of fake profiles, thereby improving the safety and trustworthiness of social media platforms.

II. COMPARATIVE STUDY

Table 2: Comparative study of previous research

Sr. No.	Author's	Title of paper	Objective of paper	Method Used	Description & Accuracy
1	**Fernando et.al[1]**	A Systematic Literature Mapping on Profile Trustworthiness in Fake News Spread	To determine the review on fake news detection using artificial intelligence	Machine learning, used to automatize the identification and combat of fake news & detection or production of Fake News.	To find the results based on the three classical profiles (persuader, clarifier, and gullible) that's considering the human and automated agents & extended the users' profile characterization and classification.
2	**Rajdavinder et al.[2]**	Detection of Fake Profiles in Online Social Networks – A Survey	Need for developing a system that is able to identifying counterfeit social sharing profile.	Graph-based features and user-based features have been mostly used.	The study suggest to improve the accuracy by using a combination of Decision Tree ,Naïve Bayes , and SVM classification algorithms.
3	**Onkar Kadam et al.[3]**	Detection of Fake Social Network Account .	To build or construct a model that can determine whether a given article is true or false.	Naive Bayes Classifier, support vector machines, and semantic analysis to identify fake news.	To determining truth and falsity in text format, as well as how and why it occurs.
4	**Michael Jonathan Ekosputra et.al.[4]**	Supervised Machine Learning Algorithms to Detect Instagram Fake Accounts.	Aims to detect Instagram fake users according to the user's profile.	supervising machine learning classification, Logistic Regression, Random Forest, the Support Vector Machine, the Artificial Neural Network, and Bernoulli Naive Bayes.	In first dataset Experiment by the using of LR,RF that will achive the accuracy 0.92. While LR,RF, 1 percent increase the accuracy 0.92 - 0.93 on the second experiment. but Bernoulli Naive Bayes had the lowest accuracy on the both experiment.
5	**Shruti Shinde, et.al.[5]**	Malicious Profile Detection on Social Media: A Survey Paper	A study of existing solutions for detecting malicious profiles on social media platforms.	Using support vector machines, random forests, JRip, and naive bayes algorithms. For the purpose of identifying fake users,	To find the highest accuracy is 95% by the help of Random Walk Method. and evaluate with some parameters such as precision, F1- Score, and recall is used to measure the result of detecting the fake profile.
6	**Ananya Bhattacharya et.al.[6]**	Application of Machine Learning Methods to Detect Fake Social Media Accounts	To evaluate a model which can deploy machine learning techniques to find rogue profiles on social media.	To applied more machine learning methods to detect authenticity on social media platform.	to find the maximum accuracy in detecting false profile's. So, prediction made 93% fake account and 96% genuine account correctly.
7	**T. Om Prathyusha et.al.[7]**	Fake Account Detection Using Machine learning .	To detect fake accounts. Gradient boosting algorithm is	Training dataset(70%) and Validation Data(10%) and Test	To analysis the dataset and used to machine learning algorithms Extreme

			used to detect fake accounts accurately.	Data(20%). And for each feature matrix fed to the classifier [LR, Random Forest, XGB, ADB, GBM]	Gradient Boosting to detect fraudulent accounts and better achieving accuracy of up to 95%.
8	**Samuel Delgado Munoz et.al.[8]**	A dataset for the detection of fake profiles on social networking services	To detect fake profiles on social media by deploying some machine learning detection methods over a novel dataset.	Decision Tree (DT), Logistic Regression, Random Forest, Multi-layer Perceptron, AdaBoost, and Gaussian Naive Bayes etc.	The best results showing Random Forest 0.96% and this algorithms obtained the optimum accuracy in combination to the best true and false prediction precision.
9	**Kumud Patel et.al.[9]**	Survey on Fake Profile Detection on Social Sites by Using Machine Learning Algorithm	Represent the review of Fake Profile Detection on Social Site by Using Machine Learning with a specific classifiers such as supervised and unsupervised machine learning.	SVM, Naïve Bays, Decision Tree, and ANNs, supervised and unsupervised machine learning.	When they truly existing data set will be give out and then determine the profile is create by a human is real or fake. And the machine learning algorithm helps in enhancing the accuracy rate of the system that is becoming 50%-96%.
10	**Pradeep Kumar Roy et.al[10]**	Fake Profile Detection on Social Networking Websites: A Comprehensive Review	We summarizes advancement of social networking's that to build a robust model to prevent and identify fake accounts on online social networking	Random forest, Naïve Bayes ,SVM, Neural Network. XGBoost, KNN etc.	Fake account identification on online social networking websites and that is classified into three categories. (1)Research using non textual (2)Research using textual feature ,(3)Research using both textual and non textual features. add new features to promote the prediction accuracy.
11	**Dr.K.Sreenivasa Rao et.al.[11]**	Detecting Fake Account on social media using machine learning Algorithms	Identifying false accounts that could mislead people and some malicious accounts are used misinformation and agenda creation.	Neural Network, Random Forest And Support Vector Machines.	To determine Whether a given account is a fake account or not depends on the dataset. Moreover, Neural Networks, Random Forest, and Support Vector Machine all shown excellent performance and the highest accuracy rate.
12	**S. P. Maniraj et.al.[12]**	Fake Account Detection using Machine Learning and Data Science.	To create new model utilized a variety of techniques to spot bogus accounts, such automated posting or comments, propagating rumors, or spam.	Data science and artificial intelligence, gradient boosting algorithm, decision tree.	The process of finding an incorrect account mainly depends on factors such as engagement rate and artificial activity to increase the efficiency of the prediction.
13	**Jyoti Singh et.al.[13]**	Detection of fake profile in social media	To detect comparable fake social network profiles that can make it simpler to connect with others in a	Machine learning, SVM, Naïve bayes (NB)	To properly identify fraudulent profiles in internet social communities as high as around 95 percent. moreover, fraudulent

			secure and effective way.		profile detection can be improved by using NLP techniques for processing texts and profiles.
14	**Kai Shu et.al.[14]**	The Role of User Profiles for Fake News Detection	To understanding and exploiting user profiles on social media and enhance the capabilities for fake news detection.	To analyze implicit and explicit features that is to obtained the information of social site including the profile features such as age, personality, location, profile image, online behavior etc.	To investigate the potential and foundation of other types of user feature and also investigate the correlations between malicious accounts and fake news and explore various user engagement behaviors such as reposts, likes, comments for fake news detection.
15	**Samala Durga Prasad Redd et.al.[15]**	Fake Profile Identification using Machine Learning	Automatic identification of fake profiles is possible and is efficient.	To use Random Forest Classifier to classify the profiles into fake or genuine classes	By using Random Forest Classifier with an extremely high level of efficiency. or real classes, which are roughly 95% and phony profile By utilizing NLP methods and neural networks, identification can be improved.
16	**Revathi.S et.al.[16]**	Profile Similarity Communication Matching Approaches for Detection of Duplicate Online social network sites	Deciding Real/Fake profile using Similarity Communication Matching algorithm.	Machine Learning, KNN, support vector machine and Node similarity communication matching.	Node Similarity Communication Matching algorithm for confirming the identity of profiles and finding out the cloned ones that investigation found that NSCM strategy accomplished an accuracy score of 93.14% in a better and decent performance.
17	**Naman Singh et.al.[17]**	Detection of Fake Profile in Online Social Networks Using Machine Learning	We set up a falsified human account as part of the task of locating, identifying, and eliminating the false accounts.	Machine learning technique Used to identify the purpose of the bots in online social media platforms.	Fake accounts easily by applying a data set and the model knows which account fake and which account is real, the model will be successfully able to differentiate a fake account created by human from a real one when the actual data set will be given to it.

III. FINDING AND DISCUSSION

In this section, we provide the process of finding a real or genuine profile to save your personal data information and also discuss the process of determining whether this profile is real or genuine. So, leveraging this framework, many machine learning strategies for categorization include Random Forest,, Support Vector Machine (SVM), Decision Tree, Naive Bayes, etc., To categorize the profiles into classes of bogus or real individuals.. In the major study, we concern all the papers that describe the processes of finding the fake and genuine, true and false profile information, etc. by utilizing a variety of machine learning methods with a specific classification that

classifies by the use of supervised machine learning, unsupervised machine learning, random forests, decision trees, naive bayes, SVM, and so on to recognize fakes and real profile information and find the better profiles. Understanding, performance, and accuracy, as compared to the researcher's paper with other people's papers and information. So, the primary purpose of fake accounts is to spread spam content, rumours, and other unauthentic messages on the platform. Hence, it is needed to filter out the fake accounts, but it has many challenges. In the past few years, researchers have applied many advanced techniques to identify fake accounts:

- Construct a false online profile with power words.
- Messages that make no sense.
- They only have a single photograph.
- They have blank profiles.
- No social networks.
- They are "renowned" or "royals."
- They are far too forward or flirtatious.8. They want your personal details.

So, in the above discussion, we learned how to use various machine learning techniques to identify the real false profile information and how many ways there are to determine whether this profile is real or not.

IV. CONCLUSION AND FUTURE WORK

Considering the survey research from this paper, the conclusion is clear: Fake accounts can be detected with the guidance of supervised machine learning. The models that have already been tested in this article are logistic regression, Bernoulli Naive Bayes, random forest, support vector machine, and artificial neural network (ANN). Any modification or addition done to features will affect the sharpness of each model. In this paper, we argue that making predictions from data is a strong task for machine learning. Yet it's crucial to keep in mind that machine learning is only as effective as the data that is used to train the algorithms and improve the analysis of data. So, the papers focus on the techniques that is performed in online social networks to identify false and legitimate accounts, and summarise the recent advancements in this field. Several of the recent methodologies are examined along with a few past techniques. The challenges and limitations of the existing approaches have also been summarised, and the dataset statistics are also compared to past techniques. For the detection of fake news, many surveys focus on specific cognitive and behavioural theories, driving more consistent work and better results. And also, understanding the roles of the profiles involved in fake news may support work exploring their behaviour and drive the development of new theories and a proposed conceptual framework that allows us to explore the main elements of this research area while providing a bigger picture of the theoretical and technical foundation. In future work, we can say that, from the above comparative study, we have observed that most work has been done by using a combination of Various algorithms for machine learning, viz., SVM, Adaboost, Random Forest, etc. Nonetheless, the accuracy in the state of the art is less than 96%. We are planning to apply the deep learning method for the identification of fake profiles to improve and increase the accuracy further.

CONFLICTS OF INTEREST

The authors declare that they have no conflicts of interest.

REFERENCES

[1] F. C. D. Da Silva, A. C. B. Garcia, and S. W. M. Siqueira, "A Systematic Literature Mapping on Profile Trustworthiness in Fake News Spread," in 2022 IEEE 25th International Conference on Computer Supported Cooperative Work in Design, CSCWD 2022, 2022, pp. 275–279. doi: 10.1109/CSCWD54268.2022.9776232.

[2] R. Singh Boparai and R. Bhatia, "Detectionof Fake Profiles in Online Social Networks-A Survey." [Online]. Available: https://ssrn.com/abstract=4159087

[3] O. Kadam and N. Surse, "Detection of Fake Social Network Account," vol. 6, pp. 2456–0774, 2021, doi: 10.51319/2456-0774.2021.4.0013.

[4] M. J. Ekosputra, A. Susanto, F. Haryanto, and D. Suhartono, "Supervised Machine Learning Algorithms to Detect Instagram Fake Accounts," in 2021 4th International Seminar on Research of Information Technology and Intelligent Systems, ISRITI 2021, 2021, pp. 396–400. doi: 10.1109/ISRITI54043.2021.9702833.

[5] S. Shinde and S. B. Mane, "Malicious Profile Detection on Social Media: A Survey Paper," in 2021 9th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions), ICRITO 2021, 2021. doi: 10.1109/ICRITO51393.2021.9596322.

[6] A. Bhattacharya, R. Bathla, A. Rana, and G. Arora, "Application of Machine Learning Techniques in Detecting Fake Profiles on Social Media," in 2021 9th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions), ICRITO 2021, 2021. doi: 10.1109/ICRITO51393.2021.9596373.

[7] T. Om Prathyusha, N. S. Kumar, and E. V. Priya, "FAKE ACCOUNT DETECTION USING MACHINE LEARNING," 2021. [Online]. Available: www.ijcrt.org

[8] S. D. Munoz and E. Paul Guillen Pinto, "A dataset for the detection of fake profiles on social networking services," in Proceedings - 2020 International Conference on Computational Science and Computational Intelligence, CSCI 2020, Dec. 2020, pp. 230–237. doi: 10.1109/CSCI51800.2020.00046. 10.1109/CSCI51800.2020.00046.

[9] Amity University, Amity University. Amity Institute of Information Technology, Institute of Electrical and Electronics Engineers. Uttar Pradesh Section, and Institute of Electrical and Electronics Engineers, ICRITO'2020 : IEEE 8th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions) : conference date: 4-5 June 2020 : conference venue: Amity University, Noida, India.

[10] C. Zhang, S. Feng, X. Wang, and Y. Wang, "ZJU-Leaper: A Benchmark Dataset for Fabric Defect Detection and a Comparative Study," IEEE Trans. Artif. Intell., vol. 1, no. 3, pp. 219–232, Dec. 2020, doi: 10.1109/TAI.2021.3057027.

[11] K. Sreenivasa Rao, S. Gutha, B. Deevena Raju, D. Rao, and D. Bdeevena Raju, "Detecting Fake Account on Social Media Using Machine Learning Algorithms System and Method For Mapping Entities Securely View Project Wireless Networks Communication View Project Detecting Fake Account On Social Media Using Machine Learning Algorithms," Int. J. Control Autom., vol. 13, no. 1s, pp. 95–100, 2020, [Online]. Available: https://www.researchgate.net/publication/340816200

[12] S. P. Maniraj, G. Harie Krishnan, T. Surya, and R. Pranav, "Fake account detection using machine learning and data science," Int. J. Innov. Technol. Explor. Eng., vol. 9, no. 1, pp. 583–585, Nov. 2019, doi: 10.35940/ijitee.A4437.119119.

[13] J. Singh and M. Z. Khan, "Issue 6 www.jetir.org (ISSN-2349-5162)," JETIR, 2019. [Online]. Available: www.jetir.org
[14] 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, ASONAM 2019, Aug. 2019, pp. 436–439. doi: 10.1145/3341161.3342927.
[15] S. Durga and P. Reddy, "Fake Profile Identification using Machine Learning," Int. Res. J. Eng. Technol., 2019, [Online]. Available: http://www.sixdegrees.com
[16] 2018 3rd International Conference on Computational Systems and Information Technology for Sustainable Solutions (CSITSS). IEEE, 2018.
[17] 2018 International Conference on Advances in Computing and Communication Engineering (ICACCE). IEEE, 2018.

ABOUT THE AUTHORS

Shamim Ahmad is M.Tech Scholar in Department of Computer Science and Engineering from Integral University, Uttar Pradesh, Lucknow, India. His researth area is machine learning. His email id is khanshamimahmad981@gmail.com

Dr. Manish Madhava Tripathi is an Associate Professor in Department of Computer Science & Engineering from Integral University, Uttar Pradesh, Lucknow.He has published many research paper in different different International journal and conferences.