

# **Detection of Fake Account on Social Media**

**Devata R. Anekar<sup>1</sup>, Arti Choudhari<sup>2</sup>, Sahil Bhosale<sup>3</sup>, Nandita Kharatmal<sup>4</sup>, Ajay Jagtap<sup>5</sup>**

Assistant Professor, Department of Information Technology<sup>1</sup>

Scholar, Department of Information Technology<sup>2,3,4,5</sup>

Sinhgad Academy of Engineering, Savitribai Phule Pune University, Pune, India

**Abstract:** *In the current generation, online social networking (OSNs) has become more popular, and social media is becoming more and more associated with these sites. They use OSN to communicate with others, share news, organize events, and run their own e-business. The strong growth of OSNs and the large number of personal information of its subscribers has led attackers, and hypocrites to steal their information, share false news, and spread malicious activities. Fake or man-made fake profiles designed to spread rumors, identity theft etc. So, in this project, we are trying to propose a discovery model, which distinguishes between fake profiles and real profiles on Twitter based on visual features such as fan counts, friends count, status calculations and more using various machine learning methods.*

**Keywords:** Fake Profile, Machine Learning, Support Vector Machine (SVM), Random Forest (RF) and KNN, Dataset.

## **I. INTRODUCTION**

Artificial intelligence can take many different meanings in different contexts, but a very brief description of Britannica, defines artificial intelligence as the 'computer power to perform tasks that are usually associated with humans. In today's world, it seems that the AI is everywhere, from the most obvious use of self-driving cars to the most obscure as the complimentary programs available on popular platforms like Netflix and Amazon. Machine learning is a subset of AI that contains any computer program that can predict without human intervention.

Identity is an article appended to an individual, separate from that person. A regular model is the name of an individual. Another model is a visa that contains the name, birth date and spot of the individual, identity, and carefully caught unique finger impression and a carefully put away and a photo of the individual. A third model is a private and public key holding fast to a Public Key Infrastructure. As a general rule, character ought to be extraordinary as in each distinguishing object should just allude to all things considered one individual. A similar individual may in any case have a few personalities, similar to an identification and a couple of keys above, or a federal retirement aide number. The genuine personality is confirmed by specialists of some country state. A modern passport is a typical example of this. Authorities guarantee that the picture, fingerprints, name, birth date etc. belong to the same person, i.e. certify the object attachment. At a social media site a user is usually identified by a profile. It typically contains a picture and name, possibly an address and birth date. The sites do not, however, rigorously check that the person with the identity alluded to in the profile really created and controls the profile. If this is not the case, somebody is using somebody else's identity. This is called false identity. One can also create profiles that can use freely invented names and other information that cannot be attached to any real person in any country.

Online Social Network, such as Facebook, Twitter, Instagram, WhatsApp and LinkedIn, have become increasingly popular over last few years, people use OSNs to keep in touch with each other's share news, organize events and even run their own e-business. The rapid growth of OSNs and the massive amount of personal data of its subscribers have attracted attackers, and imposters to steal personal data, share false news, and spread malicious activities. On the other hand researchers have started to investigate an efficient techniques to detect abnormal activities and fake accounts relying on accounts features, and classification algorithms.

## **II. RELATED WORK**

In this, paper[1] the trained model using supervised machine algorithms independently for both data sets such as fake and genuine. Ensemble classifiers have been used for the prediction more accurately as shown in figure 2. Based on research analysis, there is no such model in being used for the detection of both fake and genuine profiles on social sites. Therefore,

there are more than one machine learning algorithms are used for the detection of fake and genuine profiles. In this paper[2] they assess whether promptly accessible and designed highlights that are utilized for the effective identification, utilizing AI models, of phony personalities made by bots or PCs can be utilized to recognize counterfeit characters made by people. This is done with the expectation that comparable elements can fill in as an impetus for uncovering character double dealing by people on SMPs.

In this paper[3], a grouping algorithm has been utilized by running the Neural Network (NN) classification calculation on the choice qualities coming about because of the Support vector machine (SVM), this calculation utilizes less number of elements, while as yet having the option to accurately order around 98 percent of the records of our preparation dataset. Furthermore, we additionally approved the identification execution of our classifiers more than two different arrangements of genuine and phony records.

In this study[4], machine learning-based methods were used to detect fake accounts that could mislead people. For this purpose, the dataset generated was pre-processed and fake accounts were determined by machine learning algorithms. Decision trees, logistic regression and support vector machines algorithms are used for the detection of fake accounts. Classification performances of these methods are compared and the logistic regression proved to be more successful than the others. S. L. Bangare et al. [5-11] have worked in the brain tumor detection. N. Shelke et al [12] given LRA-DNN method. Suneet Gupta et al [13] worked for end user system. Gururaj Awate et al. [14] worked on Alzheimers Disease. P. S. Bangare et al [15-17] worked on the object detection. Kalpana Thakare et al [18-23] have worked on various machine learning algorithms. M. L. Bangare et al. [24-25] worked on the cloud platform. Rajesaheb R. Kadam et al [26] and Sachindra K. Chavan et al. [27] have discussed security issues with cloud.

### III. PROBLEM STATEMENT

The social networking sites are making our social lives better but nevertheless there are a lot of issues with using these social networking sites. The issues are privacy, online bullying, potential for misuse, trolling, etc. These are done mostly by using fake profiles.

### IV. SYSTEM ARCHITECTURE

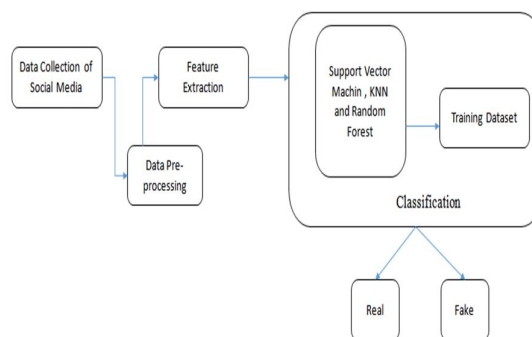


Fig. 1. System Architecture

### V. METHODOLOGY

The system is design in web development using php and python in back end. Although fake profile detection is a robust field, but it has many challenges and gaps There are a lot of existing solutions to fake profile detection but all of them have some or the other drawback. The proposed system is aiming to deliver a system which will have the highest accuracy and hence will be effective in prevention from such fake profiles by implementing and comparing different algorithms. This is done by ensemble machine learning technique which speeds up the training. In our proposed system we are aiming to design a hybrid system using support vector machine , kNN and random forest that will be able to precisely and accurately detect fake profiles in online social network. Goal of the work is to maximize the accuracy

## VI. CONCLUSION

Through utilization of different kinds of Machine Learning Algorithms, this paper is aimed to exploit different aspects of dataset which has not been deeply considered in literature and to find a good way of detection of the fake and automated accounts. In this paper we have presented a Machine Learning pipeline for detecting fake accounts in online social networks. Rather than making a prediction using one single algorithm, our system uses three different classification algorithms to determine whether or not an account in the provided dataset is a fake account or not.

## REFERENCES

- [1]. 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), 2018.
- [2]. Mohammadreza Mohammadrezaei, Mohammad Ebrahim Shiri, Amir Masoud Rahmani, "Identifying Fake Accounts on Social Networks Based on Graph Analysis and Classification Algorithms", Security and Communication Networks, vol.2018.
- [3]. Sk.Shama, K.Siva Nandini, P.Bhavya Anjali, K. Devi Manaswi "Fake Profile Identification in Online Social Networks" International Journal of Recent Technology and Engineering (IJRTE) ISSN: 2277-3878, Volume-8 Issue-4, November 2019.
- [4]. K. Ojo, A. (2019). Improved Model for Detecting Fake Profiles in Online Social Network: A Case Study of Twitter. Journal of Advances in Mathematics and Computer Science, 2019.
- [5]. S. L. Bangare, G. Pradeepini, S. T. Patil, "Implementation for brain tumor detection and three dimensional visualization model development for reconstruction", ARPN Journal of Engineering and Applied Sciences (ARPN JEAS), Vol.13, Issue.2, ISSN 1819-6608, pp.467-473. 20/1/2018 [http://www.arpnjournals.org/jeas/research\\_papers/rp\\_2018/jeas\\_0118\\_6691.pdf](http://www.arpnjournals.org/jeas/research_papers/rp_2018/jeas_0118_6691.pdf)
- [6]. S. L. Bangare, S. T. Patil et al, "Reviewing Otsu's Method for Image Thresholding." International Journal of Applied Engineering Research, ISSN 0973-4562, Volume 10, Number 9 (2015) pp. 21777-21783, © Research India Publications <https://dx.doi.org/10.37622/IJAER/10.9.2015.21777-21783>
- [7]. S. L. Bangare, G. Pradeepini, S. T. Patil, "Regenerative pixel mode and tumor locus algorithm development for brain tumor analysis: a new computational technique for precise medical imaging", International Journal of Biomedical Engineering and Technology, Inderscience, 2018, Vol.27 No.1/2. <https://www.inderscienceonline.com/doi/pdf/10.1504/IJBET.2018.093087>
- [8]. S. L. Bangare, A. R. Khare, P. S. Bangare, "Quality measurement of modularized object oriented software using metrics", ICWET '11: Proceedings of the International Conference & Workshop on Emerging Trends in Technology, February 2011, pp. 771-774. <https://doi.org/10.1145/1980022.1980190.1>
- [9]. S. L. Bangare, G. Pradeepini and S. T. Patil, "Brain tumor classification using mixed method approach," 2017 International Conference on Information Communication and Embedded Systems (ICICES), 2017, pp. 1-4, doi: 10.1109/ICICES.2017.8070748.
- [10]. S. L. Bangare, S. Prakash, K. Gulati, B. Veeru, G. Dhiman and S. Jaiswal, "The Architecture, Classification, and Unsolved Research Issues of Big Data extraction as well as decomposing the Internet of Vehicles (IoV)," 2021 6th International Conference on Signal Processing, Computing and Control (ISPPCC), 2021, pp. 566-571, doi: 10.1109/ISPPCC53510.2021.9609451.
- [11]. S. L. Bangare, G. Pradeepini, S. T. Patil et al, "Neuroendoscopy Adapter Module Development for Better Brain Tumor Image Visualization", International Journal of Electrical and Computer Engineering (IJECE) Vol. 7, No. 6, December 2017, pp. 3643-3654. <http://ijece.iaescore.com/index.php/IJECE/article/view/8733/7392>
- [12]. N. Shelke, S. Chaudhury, S. Chakrabarti, S. L. Bangare et al. "An efficient way of text-based emotion analysis from social media using LRA-DNN", Neuroscience Informatics, Volume 2, Issue 3, September 2022, 100048, ISSN 2772-5286, <https://doi.org/10.1016/j.neuri.2022.100048>.
- [13]. Suneet Gupta, Sumit Kumar, Sunil L. Bangare, Shibili Nuhmani, Arnold C. Alguno, Issah Abubakari Samori, "Homogeneous Decision Community Extraction Based on End-User Mental Behavior on Social Media",

- Computational Intelligence and Neuroscience, vol. 2022, Article ID 3490860, 9 pages, 2022. <https://doi.org/10.1155/2022/3490860>.
- [14]. Gururaj Awate, S. L. Bangare, G. Pradeepini and S. T. Patil, "Detection of Alzheimers Disease from MRI using Convolutional Neural Network with Tensorflow", arXiv, <https://doi.org/10.48550/arXiv.1806.10170>
- [15]. P. S. Bangare, S. L. Bangare, R. U. Yawle and S. T. Patil, "Detection of human feature in abandoned object with modern security alert system using Android Application," 2017 International Conference on Emerging Trends & Innovation in ICT (ICEI), 2017, pp. 139-144, doi: 10.1109/ETIICT.2017.7977025.
- [16]. P. S. Bangare and S. L. Bangare. "The Campus Navigator: An Android Mobile Application." International Journal of Advanced Research in Computer and Communication Engineering 3, no. 3 (2014): 5715-5717.
- [17]. P. S. Bangare, N. J. Uke, and S. L. Bangare, "An approach for detecting abandoned object from real time video." International Journal of Engineering Research and Applications (IJERA) 2.3 (2012): 2646-2649.
- [18]. Kalpana S. Thakare, Viraj Varale, "Prediction of Heart Disease using Machine Learning Algorithm", Bioscience Biotechnology Research Communications (Special issue) Volume 13, Issue 12, 2020 (Dec 2020 issue).
- [19]. Kalpana S. Thakare, A. M. Rajurkar, "Shot Boundary Detection of MPEG Video using Biorthogonal Wavelet Transform", International Journal of Pure and Applied Mathematics, Volume 118, No. 7, pp. 405-413, ISSN: 1311-8080 (printed version); ISSN: 1314-3395 (on-line version), url: <http://www.ijpam.eu>
- [20]. Kalpana S. Thakare, A. M. Rajurkar, R. R. Manthalkar, "Video Partitioning and Secured Key frame Extraction of MPEG Video", Proceedia Computer Science Journal, Volume 78, pp 790-798, Elsevier, 2016. Scopus DOI: <http://10.1016/j.procs.2016.02.058>, [www.sciencedirect.com/science/article/pii/S1877050916000600](http://www.sciencedirect.com/science/article/pii/S1877050916000600)
- [21]. [21] Kalpana S. Thakare, A. M. Rajurkar and R. R. Manthalkar, "Content based Video Retrieval using Latent Semantic Indexing and Color, Motion and Edge Features", International Journal of Computer Applications 54(12):42-48, September 2012, Published by Foundation of Computer Science, New York, USA. DOI: 10.5120/8621-2486
- [22]. Kalpana S. Thakare, Archana M. Rajurkar, R. R. Manthalkar, "A Comprehensive System Based on Spatiotemporal Features Such as motion, Quantized Color and Edge Features", International Journal of Wireless and Microwave Technologies (IJWMT) ISSN 1449 (Print), ISSN: 2076-9539 (Online), Vol.1, No.3, June. 2011, DOI: 10.5815 /ijwmt
- [23]. Kalpana S. Thakare, Archana M. Rajurkar, Dr. R. R. Manthalkar, "An effective CBVR system based on Motion, Quantized color and edge density features", International Journal of Computer Science & Information Technology (IJCSIT), ISSN 0975 – 3826, Vol 3, No 2, April 2011 DOI: 10.5121/ijcsit.2011.3206 78.
- [24]. M. L. Bangare, "Attribute Based Encryption And Data Integrity For Attack on Cloud Storage", Journal of Analysis and Computation (JAC), (An International Peer Reviewed Journal), [www.ijaonline.com](http://www.ijaonline.com), ISSN 0973-2861, ICASETMP-2019, pp.1-4. <http://www.ijaonline.com/wp-content/uploads/2019/07/ICASETMP67.pdf>
- [25]. M. L. Bangare, Sarang A. Joshi, "Kernel interpolation-based technique for privacy protection of pluggable data in cloud computing", International Journal of Cloud Computing, Volume 9, Issue 2-3, pp.355-374, Publisher Inderscience Publishers (IEL).
- [26]. Rajesaheb R. Kadam and Manoj L. Bangare, "A survey on security issues and solutions in live virtual machine migration", International Journal of Advance Foundation and Research in Computer (IJAFRC), (December, 2012). ISSN (2014), pp.2348-4853.
- [27]. Sachindra K. Chavan, Manoj L. Bangare, "Secure Data Storage in Cloud Service using RC5 Algorithm", International Journal of Recent Technology and Engineering (IJRTE), ISSN: 2277-3878, Volume-2, Issue-5 November 2013, pp.139-144.