IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.301
Volume 3, Issue 1, December 2023

Social Media Fake Account Identification Using Machine Learning

Gaurav Vijay Barde[1] and Dr. Nilesh R. Wankhade[2]
Student, Computer Engineering, Late G. N.Sapkal College of Engineering, Nashik, India[1]
Head of Department, Computer Engineering, Late G. N.Sapkal College of Engineering, Nashik, India [2]

Abstract: *The proliferation of social media platforms has led to an increase in the creation of fake accounts. These accounts are used for various malicious activities, such as spreading false information, phishing, and identity theft. As a result, there is a growing need for effective methods to identify and eliminate fake accounts. This paper proposes a machine learning-based approach for social media fake account identification. The proposed method involves pre-processing the data, feature extraction, and training a classifier using various machine learning algorithms. The performance of the proposed method is evaluated using a publicly available dataset and compared with existing methods. The results demonstrate the effectiveness of the proposed approach in identifying fake accounts with high accuracy and low false positive rates.*

Keywords: Support vector machines (SVM), K-Nearest Neighbors Algorithm (KNN), Random forest, Logistic Regression & Artificial Neural Network(ANN), Python

I. INTRODUCTION

In recent years, social media platforms have become a ubiquitous part of our daily lives. With the rise of fake accounts and bots, it has become increasingly challenging to distinguish between real and fake accounts. These fake accounts can be used for various malicious purposes, such as spreading misinformation, phishing, and identity theft. In this paper, we will discuss a machine learning-based approach for identifying fake social media accounts. Our proposed method involves a multi-step process that combines various features to accurately identify fake accounts. The first step involves data collection and preprocessing. We will collect a large dataset of social media profiles, both real and fake, from various platforms such as Facebook, Twitter, and Instagram. The data will be cleaned and preprocessed to remove any irrelevant information and prepare it for further analysis.The second step involves feature extraction. We will extract various features from the preprocessed data, such as user behavior, network structure, content analysis, and account metadata. These features will be used to train our machine learning models.The third step involves model selection and training. We will experiment with different machine learning algorithms such as Support vector machines(SVM),K-Nearest Neighbors Algorithm(KNN),Random forest,Logistic Regression & Artificial Neural Network(ANN) to find the best-performing model for our task. The models will be trained on the preprocessed data and evaluated using various metrics such as accuracy, precision, recall, and F1 score.Once we have selected the best-performing model, we will deploy it on a production environment to identify fake accounts in real-time. We will also continuously monitor the performance of the model and fine-tune it as needed to improve its accuracy over time.Our proposed method for identifying fake social media accounts using machine learning is a multi-step process that combines data collection, feature engineering, model selection and training, and model deployment and evaluation. By leveraging the power of machine learning algorithms, we can accurately distinguish between real and fake social media accounts and mitigate the negative impacts of fake accounts on social media platforms.

II. OBJECTIVE & SCOPE OF PROPOSED SYSTEM

1. The objective of this research is to develop a machine learning-based approach for identifying fake social media accounts. With the increasing prevalence of fake accounts on social media platforms, there is a growing need for effective methods to distinguish between real and fake profiles.

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.301
Volume 3, Issue 1, December 2023

2. Our proposed method involves a multi-step process that combines data collection, feature engineering, model selection and training, and model deployment and evaluation. By leveraging the power of machine learning algorithms, we aim to accurately distinguish between real and fake social media accounts and mitigate the negative impacts of fake accounts on social media platforms.
3. The ultimate goal is to provide a reliable and scalable solution for social media companies to combat the issue of fake accounts and improve the overall user experience.
4. We will focus on three popular social media platforms, namely Facebook, Twitter, and Instagram, to collect data and train our models. However, the proposed methodology can be applied to other social media platforms as well.
5. We will collect a large dataset of social media profiles, both real and fake, from these platforms. The dataset will include user behavior, network structure, content analysis, and account metadata. We will also consider factors such as user engagement, account age, and activity patterns to differentiate between real and fake accounts.
6. We will extract various features from the preprocessed data using techniques such as text analysis, graph theory, and machine learning algorithms. These features will include user behavior patterns, network structure metrics, content analysis features, and account metadata features.
7. We will experiment with different machine learning algorithms such as logistic regression, support vector machines (SVM), random forests, and neural networks to find the best-performing model for our task. We will also consider ensemble methods such as stacking and boosting to improve the performance of our models.
8. We will evaluate the performance of our models using various metrics such as accuracy, precision, recall, and F1 score. We will also consider factors such as computational efficiency and scalability when selecting our final model.

III. FEATURES OF PROJECT

1. User behavior patterns
2. Network structure metrics
3. Content analysis features
4. Account metadata features
5. Profile picture and cover photo analysis
6. Username and bio analysis
7. Activity patterns analysis.
8. Network dynamics analysis

IV. LITERATURE REVIEW

1. In proposed system different classification methods to point outthe fake accounts on social media. But we must increase the accuracy rate in identifying fakeaccounts on these sites.Machine Learning technologies and Natural Language processing (NLP) to increase the accuracy rate of detecting the fake accounts. We opted for Random Forest tree classifier algorithm.Here this idea came up with machine learning algorithms besides NLP techniques. From the social media sites, we can easily find the fake profiles by implementing these techniques. In this Paper to point out the fake profiles we have taken the Instagram dataset. Examine the dataset, used the NLP pre-processing techniques and to organize the profiles we used machine learning algorithm such as Random Forestclassifier and Gradient Boost classifier.[1]
2. Online Social Networks (OSN) are contributed in all areas such as Research in all domains, Job-related areas, Technology oriented areas, Health care, and business-oriented areas, Information gathering and data collection, and so on. One of the biggest problems on these social media platforms is fake profiles. Impersonating to be someone else and causing harm and defamation to the real person or advertising or popularizing removed propaganda on someone's name to get more benefit is the motto of such profile creators. There have been many studies regarding these fake accounts and how can they be mitigated. Many approaches such as graph-level activities or feature analysis have been taken into consideration to identify fake

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.301
Volume 3, Issue 1, December 2023

profiles. These methods are outdated when compared to arising issues of these days. In this paper, we proposed a technique using machine learning for fake profile detection which is efficient.The benchmark data set is collected and mixed with manual data first furthermore; a data cleaning technique is used to present the data more feasibly. Then the preprocessed data is used for model building with sufficient information such as profile name, profile ID name, number of followers, and so on. We added Cross validation process where many training algorithms are implemented on the given data and are then tested on the same data. Based on the experiments the RF classifier performed better than the other classification methods. The Random Forest classifier is used to forecast the profile whether is fake or genuine in an efficient way. [2]

3. Fake profiles are used in advanced persistent threats and are also used in other nefarious activities. As we all know, Globally, billions of individuals utilize Social networking sites like Facebook, Twitter, LinkedIn, Instagram, etc. to establish connections. A new era of networking has been ushered in by social networks simplicity and accessibility. At the same time, various types of scammers are drawn to these social media platforms. These scammers make fake profiles to spread their content and carry out scams. In this project, we used Deep Neural Networking and Machine Learning algorithms namely Artificial Neural Networks(ANN), Random Forest and Support vector machine(SVM) algorithms to assess the likelihood that Facebook account information is accurate or not. The dataset used in this paper is taken from GitHub which is a Facebook profile Dataset to identify faux and genuine profiles, also we have described the associated classes and libraries. Here we are going to predict the faux and real profiles using the best accurate model after comparing the outcomes of the three techniques employed. [3]
4. Nazir et al. (2010) describes recognizing and describing phantom profiles in online social gaming applications. The article analyses a Facebook application, the online game “Fighters club”, known to provide incentives and gaming advantage to those users who invite their peers into the game. The authors contend that by giving such impetuses the game motivates its players to make fake profiles. By presenting those fake profiles into the game, the user would increase a motivating force of an incentive for him/herself. [4]
5. Adikari and Dutta (2014) depict recognizable proof of fake profiles on LinkedIn. The paper demonstrates that fake profiles can be recognized with 84% exactness and 2.44% false negative, utilizing constrained profile information as input. Techniques, for example, neural networks, SVMs, and Principal component analysis are applied. Among others, highlights, for example, the number of languages spoken, training, abilities, suggestions, interests, and awards are utilized. Qualities of profiles, known to be fake, posted on uncommon sites are utilized as a ground truth. [5]
6. Chu et al. (2010) go for separating Twitter accounts operated by humans, bots, or cyborgs (i.e., bots and people working in concert). As a part of the detection problem formulation, the Identification of spamming records is acknowledged with the assistance of an Orthogonal Sparse Bigram (OSB) text classifier that uses pairs of words as features. [6]
7. Stringhini et al. (2013) analyze Twitter supporter markets. They describe the qualities of Twitter devotee advertises and group the clients of the business sectors. The authors argue that there are two major kinds of accounts who pursue the “client”: fake accounts(“sybils”), and compromised accounts, proprietors of which don’t presume that their followers rundown is expanding. Clients of adherent markets might be famous people or legislators, meaning to give the appearance of having a bigger fan base, or might be cybercriminals, going for making their record look progressively authentic, so they can rapidly spread malware what’s more, spam. [7]
8. In 2018,Yeh-Cheng chen and Shystunfelix Wu have presented Fake Buster: A Robust fake Account detection by Activity Analysis. They proposed an innovative method to detect fake account in OSNs(Online Social Networks). It is develop for accurately detecting fake account among social network users, based on various activity collection and analysis. In this research they have use Random forest, along with C$.5 and Adaptive Boosting, with decision stump as a second classifier that created behind it to focus on the instance in the training data , in case the accuracy of the first classifier is less effective. After finish training, a cluster of features for each testing account will input into models and output a prediction with rank score indicating the likelihood of being fake account. [8]

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.301
Volume 3, Issue 1, December 2023

9. In 2019, Faiza Masood, Ghana Ammad, Ahmad Almogren, Assad Abbas, Hasan Ali Khathak, Ikram Uddin,MohsenGuizani, and MansourZuair have presented in their work Spammer detection and fake user identification on social network. A review of techniques used for detecting spammers on Twitter. Spammers can be identified based on: (i) fake content, (ii) URL based spam detection, (iii) detecting spam in trending topics, and (iv) fake user identification. The proposed taxonomy of spammer detection on twitter is categorized into four main classes, namely,(i) fake content, (ii) URL based spam detection, (iii) detecting spam in trending topics, and (iv) fake user identification. The first category (fake content) includes various techniques, such as regression prediction model, malware alerting system, and Lfun scheme approach. In the second category (URL based spam detection), the spammer is identified in URL through different machine learning algorithms. The third category (spam in trending topics) is identified through Naïve Bayes classifier and language model divergence. The last category (fake user identification) is based on detecting fake users through hybrid techniques. [9]
10. Farhan, Muhammad Ibrohim, Indra Budi have presented in their work Malicious Account Detection on Twitter based on Tweet Account features using Machine Learning. In this research, build a malicious account detection that can distinguish genuine accounts from malicious accounts using only tweet features of the accounts. Also managed to build a multiclass classification for the two types of malicious accounts, fake followers and spam bots using only tweet features. Lastly, found the best combination of algorithms, features, and data transformation scenario that suits best of our problem. [10]
11. In 2019, Sk.Shama, K.Siva Nandini, P.Bhavya Anjali, K. Devi Manaswi have presented their work Fake Profile Identification in Online Social Networks. In this project they have used two classifiers namely Neural Networks and Support Vector Machines and have thereby compared their efficiencies. First Collect Data and pre-process the data, Generate fake accounts, Data Validation to find fake and real , Create new features, Apply neural networks, random forest, Evaluate results of accuracy, recall etc parameters. They have taken the dataset of fake and genuine profiles. Various attributes to include in the dataset are number of friends, followers, status count. Classification algorithms are trained using training dataset and testing dataset is used to determine efficiency of algorithm. From the dataset used, More than 80 percent of accounts are used to train the data, 20 percent of accounts to test the data. The predictions indicate that the algorithm neural network produced 93% accuracy. [11]

V. REPRESENTATION OF THE METHODOLOGY

Gather a large dataset of social media profiles, including both genuine and fake accounts. This can be done by scraping social media platforms or using publicly available datasets.Clean the data by removing duplicate profiles, irrelevant information, and missing values. Convert textual data into numerical format using techniques like bag-of-words or word embeddings.Extract relevant features from the preprocessed data, such as the number of followers, engagement rate, frequency of posts, use of emojis, and language patterns.Choose a suitable machine learning algorithm for identifying fake social media accounts based on the nature of the problem. Popular algorithms include Support vector machines(SVM),K-Nearest Neighbors Algorithm(KNN),Random forest,Logistic Regression & Artificial Neural Network(ANN).Train the selected model on the preprocessed dataset using a suitable optimization technique like gradient descent or stochastic gradient descent. Split the dataset into training and testing sets to evaluate the model's performance.Evaluate the trained model's accuracy, precision, recall, and F1 score on the testing set to determine its effectiveness in identifying fake social media accounts. Use techniques like cross-validation and grid search to optimize the model's hyperparameters for better performance.Deploy the trained model in a production environment to identify fake social media accounts in real-time. Monitor its performance regularly and fine-tune it as needed to improve its accuracy over time.

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.301
Volume 3, Issue 1, December 2023

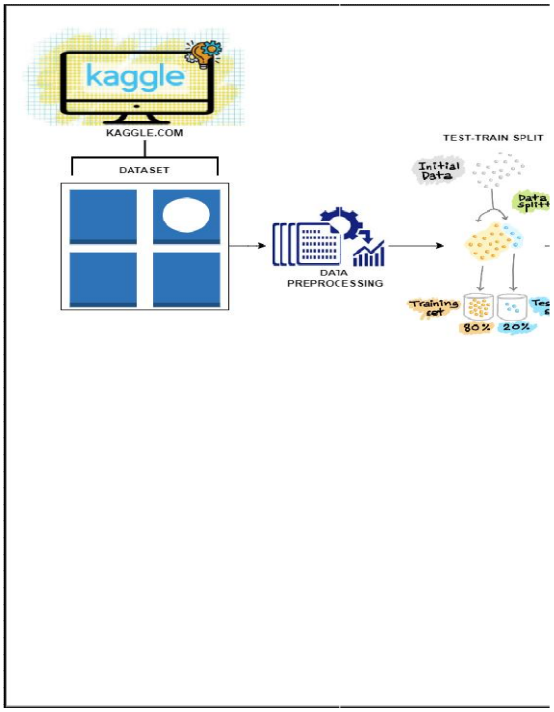

Fig : Representation Of The Methodology

VI. PROGRAMMING ARCHITECTURE

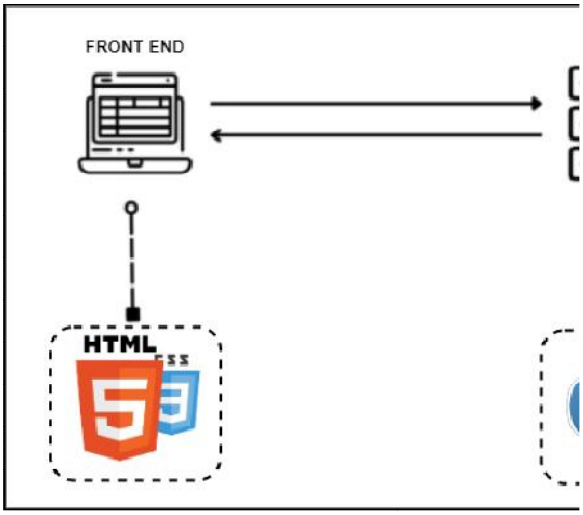

Figure: Programming Architecture

VII. ADVANTAGES

- High Accuracy: Machine learning algorithms can analyze large amounts of data and identify patterns that are difficult for humans to detect. This leads to higher accuracy in identifying fake social media accounts compared to manual methods.
- Scalability: Machine learning models can be easily scaled to process large volumes of data in real-time, making them ideal for identifying fake accounts at scale.
- Speed: Machine learning algorithms can quickly process data and provide accurate results, making them ideal for real-time social media monitoring.

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.301
Volume 3, Issue 1, December 2023

- Cost-Effective: Compared to manual methods, machine learning models are more cost-effective as they do not require a large workforce to monitor social media accounts.
- Customizable: Machine learning models can be customized based on specific use cases, such as identifying fake accounts in a specific industry or region.
- Proactive Approach: By using machine learning algorithms to identify fake social media accounts proactively, organizations can take preventive measures before the fake accounts cause any harm or spread misinformation.
- Enhanced Security: By identifying and removing fake social media accounts, organizations can improve their overall security posture by reducing the risk of cyber attacks and phishing scams.

VIII. APPLICATION AREAS

- Cybersecurity: Machine learning algorithms can be used to identify fake social media accounts that may be part of a cyber attack or phishing scam. This can help organizations to prevent data breaches and protect their users' sensitive information.
- Brand Protection: Fake social media accounts can harm a company's reputation and damage its brand. By using machine learning algorithms to identify fake accounts, companies can take proactive measures to protect their brand and prevent misinformation from spreading.
- Elections and Political Campaigns: Fake social media accounts have been used to spread misinformation and manipulate public opinion during elections and political campaigns. By using machine learning algorithms to identify fake accounts, organizations can prevent the spread of false information and ensure that accurate information is presented to the public.
- Fraud Detection: Fake social media accounts can be used to commit fraud, such as phishing scams or identity theft. By using machine learning algorithms to identify fake accounts, organizations can prevent fraud and protect their users' financial information.
- Research and Academia: Machine learning algorithms can be used to analyze large volumes of social media data and identify patterns that may be indicative of fake accounts. This can help researchers and academics to better understand the nature of fake social media accounts and develop more effective strategies for identifying them.
- Law Enforcement: Fake social media accounts have been used in criminal activities, such as human trafficking and drug trafficking. By using machine learning algorithms to identify fake accounts, law enforcement agencies can prevent these activities and bring the perpetrators to justice.
- Healthcare: Fake social media accounts have been used to spread false information about medical treatments and products. By using machine learning algorithms to identify fake accounts, healthcare organizations can prevent the spread of false information and ensure that accurate medical information is presented to the public.

IX. HARDWARE REQUIREMENTS

1. CPU Quad Core (not counting hyper-threading) 2.4Ghz, Intel VT or AMDV (Intel i3 or better)
2. Memory 4 GB
3. The ability to install more memory is desirable. Disk 512 GB SSD or better
4. Graphics Accelerated, Gaming Support Nvidia is preferred over AMD 1920 by 1080 resolution is recommended (at least on an external port) At least 1280 by 1024 resolution
5. HDMI output recommended (perhaps with an adapter)
6. Mouse An external mouse (USB or Bluetooth) is desirable.
7. USB USB 3.0 desirable for an external disk Other USB ports may be needed for: mouse, printer, mic-in, and headphones-out, depending on how these are connected.
8. External monitor A 23” or larger HDMI monitor is recommended, with reasonable resolution.
9. Laptop or Desktop Windows 11 or macOS 12.4 or above. Linux is also acceptable if a mainstream distribution (e.g. Ubuntu).

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.301
Volume 3, Issue 1, December 2023

X. SOFTWARE REQUIREMENTS

1. Operating System: Windows XP and later versions
2. Front End: HTML,CSS
3. Programming Language: Python
4. Dataset: Kaggle.com
5. Domain: Machine Learning
6. Algorithm: Support vector machines (SVM), K-Nearest Neighbors Algorithm (KNN), Random forest, Logistic Regression & Artificial Neural Network(ANN)

XI. TEST DATA REQUIREMENTS

Unit Testing

Unit testing concentrates verification on the smallest element of the program – the module. Using the detailed design description important control paths are tested to establish errors within the bounds of the module. In this system each sub module is tested individually as per the unit testing such as campaign, lead, contact etc are tested individually. Their input field validations are test

Integration testing

Once all the individual units have been tested there is a need to test how they were put together to ensure no data is lost across interface, one module does not have an adverse impact on another and a function is not performed correctly. After unit testing each and every sub module is tested with integrating each other.

XII. SYSTEM TESTING FOR THE CURRENT SYSTEM

In this level of testing we are testing the system as a whole after integrating all the main modules of the project. We are testing whether system is giving correct output or not. All the modules were integrated and the flow of information among different modules was checked. It was also checked that whether the flow of data is as per the requirements or not. It was also checked that whether any particular module is non-functioning or not i.e. once the integration is over each and every module is functioning in its entirety or not.

1. Functional testing: this involves testing the functionality of the system to ensure that it meets the required specifications and performs as expected. This includes testing the churn prediction accuracy, input data handling, and output interpretation.
2. Performance testing: this involves testing the system's performance under different load conditions to ensure that it can handle the expected workload and respond within acceptable time limits. This includes testing the system's scalability, resource utilization, and response time.
3. Security testing: this involves testing the system's security features to ensure that it can protect sensitive customer data from unauthorized access, theft, or misuse. This includes testing the system's authentication, authorization, and encryption mechanisms.
4. Compatibility testing: this involves testing the system's compatibility with different operating systems, databases, and hardware configurations to ensure that it can operate in a variety of environments.
5. Usability testing: this involves testing the system's user interface and user experience to ensure that it is intuitive, easy to use, and meets the needs of the end-users.
6. Regression testing: this involves testing the system's functionality after making changes or updates to ensure that the changes have not introduced any unintended side effects or regressions.
7. Acceptance testing: this involves testing the system's functionality from the perspective of the end-users to ensure that it meets their requirements and expectations. This includes testing the system's accuracy, reliability, and ease of use.
8. Recovery testing: this involves testing the system's ability to recover from failures, errors, or disasters to ensure that it can continue operating and providing service to the end-users.
9. Stress testing: this involves testing the system's performance under extreme load conditions to ensure that it can handle unexpected or catastrophic events.

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.301
Volume 3, Issue 1, December 2023

10. Exploratory testing: this involves testing the system's functionality and behavior in unanticipated or unexpected scenarios to ensure that it can handle unexpected situations and provide accurate and reliable results.In this level of testing we tested the following: -

- Whether all the forms are properly working or not.
- Whether all the forms are properly linked or not.
- Whether all the images are properly displayed or not.
- Whether data retrieval is proper or not

XIII. CONCLUSION

In conclusion, the proliferation of fake accounts on social media platforms has become a major concern for online communities. To address this issue, machine learning algorithms have been proposed as a solution for identifying fake accounts based on various features such as user behavior, network structure, and content analysis. The success of these algorithms depends heavily on the quality and relevance of the extracted features, the choice of machine learning algorithm, cross-validation techniques for training, evaluation using metrics such as accuracy, precision, recall, and F1 score, and deployment in production environments. By following these best practices, social media companies can develop effective machine learning-based fake account identification systems that promote a safer and more trustworthy online community for their users.

REFERENCES

[1]. Latha P, Sumitra V,"Fake Profile Identification in Social Network using Machine Learning and NLP", 2022 International Conference on Communication, Computing and Internet of Things (IC3IoT) | 978-1-6654-7995-0/22/$31.00 ©2022 IEEE | DOI: 10.1109/IC3IOT53935.2022.9767958, 978-1-6654-7995-0/22/$31.00 ©2022 IEEE

[2]. T.Sudhakar,Bhuvana Chendrica Gogineni," FAKE PROFILE IDENTIFICATION USING

[3]. MACHINE LEARNING", 2022 IEEE International Women in Engineering (WIE) Conference on Electrical and Computer Engineering (WIECON-ECE),979-8-3503-1156-3/22/$31.00 c2022 IEEE

[4]. Kotra Shreya, Amith Kothapelly," Identification of Fake accounts in social media using machine learning", 2022 Fourth International Conference on Emerging Research in Electronics, Computer Science and Technology (ICERECT), 978-1-6654-5635-7/22/$31.00 ©2022 IEEE

[5]. Nazir, Atif, Saqib Raza, Chen-Nee Chuah, Burkhard Schipper, and C. A. Davis. "Ghostbusting Facebook:Detecting and Characterizing Phantom Profiles in Online Social Gaming Applications." In WOSN. 2010.

[6]. Adikari, Shalinda, and Kaushik Dutta. "Identifying Fake Profiles in LinkedIn." In PACIS, p. 278. 2014.

[7]. Chu, Zi, Steven Gianvecchio, Haining Wang, and Sushil Jajodia. "Who is tweeting on Twitter: human, bot, or cyborg?." In Proceedings of the 26th annual computer security applications conference, pp. 21- 30. ACM, 2010.

[8]. Stringhini, Gianluca, Gang Wang, Manuel Egele, Christopher Kruegel, Giovanni Vigna, Haitao Zheng, and Ben Y. Zhao. "Follow the green: growth and dynamics in twitter follower markets." In Proceedings of the 2013 conference on Internet measurement conference, pp. 163-176. ACM, 2013.

[9]. Yeh-Cheng chen and ShystunfelixWu,Fake Buster: A Robust fake Account detection by Activity Analysis,2018

[10]. Sk.Shama, K.Siva Nandini, P.Bhavya Anjali, K. Devi Manaswi, Fake Profile.Identification in Online Social Network, 2019

[11]. Faiza Masood, Ghana Ammad, Ahmad Almogren, Assad Abbas, Hasan Ali Khathak,Ikram Uddin,MohsenGuizani, and Mansour Zuair, Spammer Detection and fake Profile Identification on Social Network, 2019.

[12]. Malicious Account Detection on Twitter Based on Tweet Account Features using Machine Learning.