IJITCE

ISSN 2347- 3657

International Journal of Information Technology & Computer Engineering

Email : ijitce.editor@gmail.com or editor@ijitce.com

International Journal of Information Technology & Computer Engineering
ISSN 2347–3657
Volume 13, Issue 2, 2025

FAKE PROFILES IDENTIFICATION IN ONLINE SOCIAL NETWORKS USING MACHINE LEARNING AND NLP

[1]Mrs. Ch. Deepika Asst.Professor, [2]G. GUNA SRI, [3]V. SUPRIYA, [4]N. SUBHASHINI, [5]K. NEHA RANI
EMAIL: deepika3062@gmail.com
Vijaya Institute of Technology for Women
(Affiliated to J.N.T.U Kakinada, Approved by A.I.C.T.E, New Delhi)
DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING

ABSTRACT

Online Social Networks (OSNs) have become an integral part of modern communication, connecting millions of users worldwide. However, the proliferation of fake profiles poses significant security and privacy threats, including misinformation, fraud, and cyber harassment. Detecting and mitigating these fraudulent accounts is essential to maintaining the integrity of online platforms. Traditional rule-based approaches struggle to adapt to the evolving tactics of malicious actors, necessitating more robust and intelligent solutions. Machine Learning (ML) and Natural Language Processing (NLP) offer promising avenues for accurately identifying fake profiles by analysing behavioural and textual features. This study explores the application of ML and NLP techniques for the detection of fake profiles in OSNs. A comprehensive dataset of both genuine and fake accounts is utilized to extract key features such as profile attributes, activity patterns, and linguistic characteristics of user generated content. Various supervised and unsupervised ML models, including decision trees, support vector machines, deep learning networks, and ensemble methods, are evaluated for their effectiveness in identifying fraudulent accounts. Additionally, NLP-based techniques, such as sentiment analysis and text embedding models, are leveraged to analyse user generated content for signs of deception and unnatural linguistic patterns.

KEYWORDS: Online Social Network, Machine Learning, Natural Language Processing, Support Vector Machine, Navie Bayes, Random Forest.

1. INTRODUCTION

Online Social Networks (OSNs) have revolutionized digital communication by enabling people to connect, share information, and build relationships on a global scale. Platforms such as Facebook, Twitter, Instagram, and LinkedIn have amassed billions of users, fostering interactions that span personal, professional, and commercial domains. While OSNs offer immense benefits, they also pose significant challenges, particularly in terms of security and privacy. One of the most pressing concerns is the widespread presence of fake profiles, which are often used for deceptive purposes such as identity theft, phishing, cyberbullying, misinformation dissemination, and fraudulent activities. The ability to detect and mitigate fake profiles is

ISSN 2347–3657
Volume 13, Issue 2, 2025

crucial to ensuring the integrity of online communities and protecting users from malicious actors.

1.1 The Growing Threat of Fake Profiles

Fake profiles on OSNs are created for various deceptive purposes, including identity theft, phishing, misinformation campaigns, and fraud. They often mimic real user behaviour to avoid detection, making traditional detection methods like blacklisting and manual moderation insufficient. As OSNs grow, the volume of user accounts makes manual detection impractical, and fake profiles can blend in seamlessly.

1.2 Machine Learning and Natural Language Processing in Fake Profile Detection

ML and NLP are increasingly used to detect fake profiles. ML models can identify patterns in user behaviour and profile attributes, while NLP focuses on analysing text content for inconsistencies and unnatural language patterns. Supervised ML models use labelled data to classify profiles, while unsupervised approaches like anomaly detection help identify unknown fraudulent behaviours. NLP techniques, including sentiment analysis and entity recognition, further enhance fake profile detection by examining linguistic features of user generated content.

1.3 Challenges in Fake Profile Detection

Detecting fake profiles is complicated by the adaptability of fraudsters, who use tactics such as automated bots and strategically

2. LITERATURE REVIEW

The literature review chapter provides an in-depth analysis of previous research on fake profile detection, machine learning applications in cybersecurity, and the use of NLP for text-based fraud detection. It systematically examines academic papers, industry reports, and case studies that have explored various techniques for identifying fraudulent accounts. The goal of this chapter is to assess the strengths and limitations of existing approaches while identifying gaps that this research aims to address.

The literature review is divided into multiple subsections to cover different aspects of fake profile detection. One section focuses on traditional rule-based and heuristic approaches, explaining how early detection systems relied on predefined conditions such as profile completeness, friend network density, and posting frequency. The limitations of these methods, particularly their inability to adapt to evolving fraudulent tactics, are discussed.

EXISTING METHODOLIES:

Fake profile detection has been a persistent challenge for online social networks. Over the years, various

ISSN 2347–3657
Volume 13, Issue 2, 2025

approaches have been implemented to mitigate the risks posed by fraudulent accounts. The existing systems primarily rely on rule-based methods, behavioural analysis, network graph analysis, and machine learning-based classification. While these methods have improved the ability to detect and remove fake profiles, fraudsters continuously evolve their tactics to bypass detection mechanisms. The limitations of existing systems highlight the need for more advanced methodologies that leverage artificial intelligence and natural language processing.

Most social media platforms have implemented automated systems that analyze user activity, profile attributes, and interaction patterns to detect anomalies. These systems flag suspicious accounts based on predefined rules, historical data, or behavioural inconsistencies. However, fake profiles have become more sophisticated, using techniques such as automated bots, deepfake technology, and coordinated social engineering tactics. While the existing systems can detect simple fake profiles, they often struggle with well-disguised fraudulent accounts that mimic real users.

3. PROPOSED SYSTEM:

The increasing prevalence of fake profiles on social media has necessitated the development of a more robust and intelligent system for their identification and elimination. The proposed system leverages machine learning and natural language processing techniques to detect fraudulent accounts with a high degree of accuracy. Unlike traditional rule-based approaches, which rely on predefined heuristics, this system dynamically learns patterns and behaviours associated with fake profiles. The integration of artificial intelligence ensures adaptability to evolving deception strategies employed by malicious actors. The proposed system focuses on automating the detection process, improving accuracy, enhancing scalability, and ensuring compliance with ethical and legal standards.

3.1 System Architecture

The system follows a modular architecture consisting of several interconnected components, each performing specific functions. The main modules include data collection, data preprocessing, feature extraction, model training, fake profile classification, and user interface. The architecture is designed to handle real-time and batch processing scenarios, allowing social media platforms and cybersecurity agencies to detect fraudulent accounts efficiently.

International Journal of Information Technology & Computer Engineering
ISSN 2347–3657
Volume 13, Issue 2, 2025

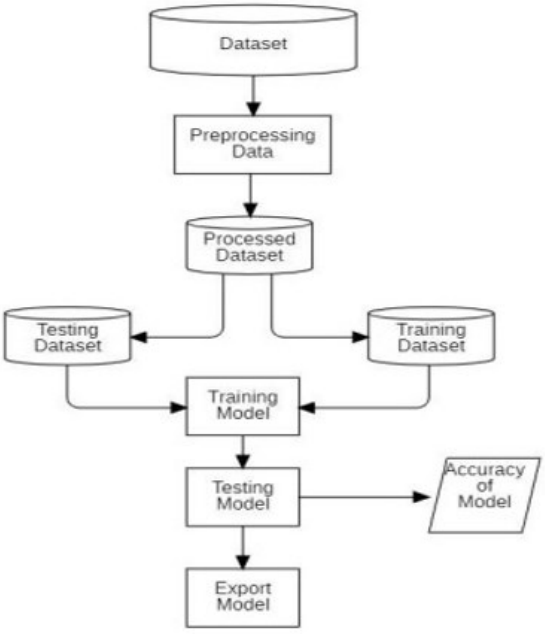

Fig 1: Architecture

4. RESULTS

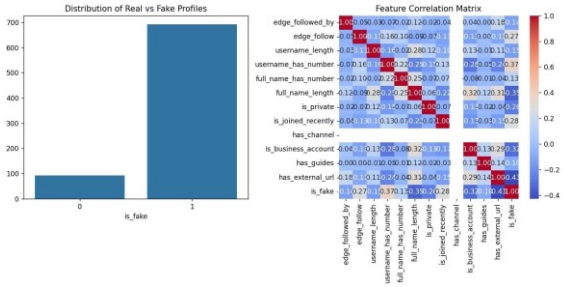

Fig 2: Correlation of Features

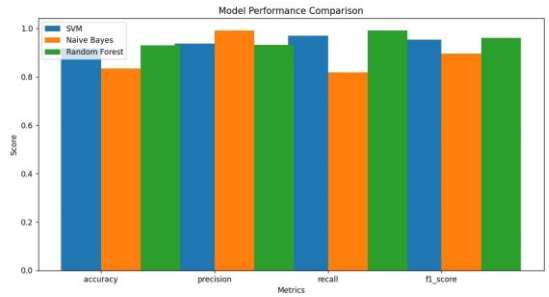

Fig 3: Model Performance Graph

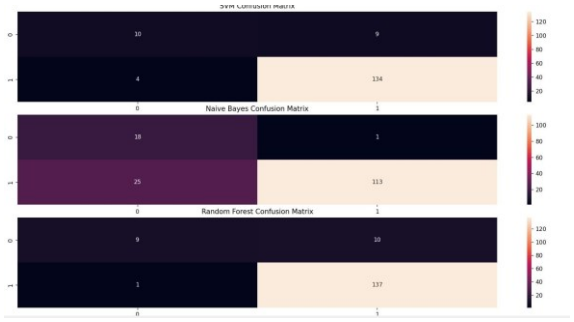

Fig 4: Confusion Matrix

CONCLUSION

The proposed system addresses the growing issue of fake profiles on social media, which pose threats like misinformation, fraud, and identity theft. It utilizes advanced machine learning (ML) and natural language processing (NLP) techniques to detect fraudulent accounts with high accuracy and efficiency administrators to take action quickly.

REFERENCE

1. Aggarwal, C. C. (2011). Social network data analytics. Springer.
2. AlZubi, A. A., Manickam, S., & Alzubi, O. (2019). Fake profile detection using supervised machine learning algorithms. Journal of Computer Science, 15(4), 456-467.
3. Benevenuto, F., Magno, G., Rodrigues, T., & Almeida, V. (2010). Detecting spammers on Twitter. Proceedings of the 7th Annual Collaboration, Electronic

Messaging, AntiAbuse and Spam Conference, 12-18.

4. Cresci, S., Lillo, F., Regoli, D., Tardelli, S., & Tesconi, M. (2019). Cashtag piggybacking: uncovering spam and bot activity in stock microblogs on Twitter. ACM Transactions on the Web, 13(2), 1-27.

5. Ferrara, E., Varol, O., Davis, C., Menczer, F., & Flammini, A. (2016). The rise of social bots. Communications of the ACM, 59(7), 96-104.

6. Kudugunta, S., & Ferrara, E. (2018). Deep neural networks for bot detection. Information Sciences, 467, 312-322.

7. Kumar, S., Spezzano, F., Subrahmanian, V., & Faloutsos, C. (2017). Edge weight prediction in weighted signed networks. IEEE Transactions on Knowledge and Data Engineering, 29(6), 1325-1338.

8. Santia, G. C., & Williams, J. R. (2018). Detecting malicious social bots: a discriminative approach. Proceedings of the 2018 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining, 1-8.

9. Varol, O., Ferrara, E., Davis, C. A., Menczer, F., & Flammini, A. (2017). Online humanbot interactions: Detection, estimation, and characterization. Proceedings of the 11th International AAAI Conference on Web and Social Media, 280-289.