Malware Net AI: Autoencoder-Based Malware Anomaly Detection System

[1]Mohammed Asra Anjum, [2]Mrs. Tanneru Venkata Lavanya
[1]M.Tech Scholar, Dept. of CSE (AI&ML), Malla Reddy Technical Campus ,Malla Reddy Vishwavidyapeeth,Maisammaguda, Hyderabad, Telangana 500100, India, asraanjum7654@gmail.com
[2]Assistant Professor, Dept. of CSE , Malla Reddy Technical Campus,Malla Reddy Vishwavidyapeeth, Maisammaguda, Hyderabad, Telangana 500100, India, vlavanyat91@gmail.com

Article Info

Received: 23-03-2026 Revised: 02-04-2026 Accepted: 10-04-2026 Published: 21-04-2026

ABSTRACT

Social media platforms have become an essential medium for communication, business, and information sharing. However, the increasing growth of these platforms has also given rise to fake accounts that pose threats such as misinformation, cyber fraud, spam, and identity impersonation. Detecting and preventing fake accounts is therefore a critical challenge to ensure the security, reliability, and trustworthiness of online communities. This project focuses on the detection of fake accounts on social media using machine learning techniques. By analyzing user profile features, activity patterns, content characteristics, and network behavior, the system can differentiate between genuine and fake users. The study explores both traditional rule-based methods and modern algorithms such as Decision Trees, Random Forests, and XGBoost, which improve detection accuracy. The proposed approach emphasizes building an intelligent model that learns from real-world datasets and adapts to evolving patterns of fake account creation. The outcome of this project will help in reducing the spread of spam, safeguarding users from fraudulent activities, and enhancing the overall trust of social media platforms. This work not only contributes to the technical understanding of fake account detection but also highlights the importance of ethical considerations, user privacy, and scalability in real-world applications. The project aims to demonstrate that advanced machine learning models can provide effective and reliable solutions to one of the most pressing issues in today's digital era.

INTRODUCTION

In recent years, social media has become an integral part of human communication, business, and entertainment. Platforms like Facebook, Twitter, Instagram, and LinkedIn connect millions of users worldwide, enabling them to share ideas, build communities, and stay updated with current events. However, the rapid growth of these platforms has also led to the rise of fake accounts, which represent a significant challenge to both users and organizations. A fake account can be defined as an online profile that does not represent a real person or misrepresents its identity. In many cases, fake accounts are used to promote false products, increase follower counts artificially, or influence political discussions. Such activities threaten the trustworthiness, safety, and credibility of social media platforms. Detecting fake accounts is not a simple task, as they are often designed to imitate real users. While some may have incomplete profiles or suspicious usernames, advanced fake accounts make use of stolen images, realistic posts, and natural-looking interactions. Despite these attempts to appear genuine, fake accounts can often be identified by analyzing their unusual behavior. For example, they may send a large number of friend requests in a short time, post repetitive or irrelevant content, engage in abnormal liking/commenting patterns, or form suspicious network connections. To address this problem, researchers and developers have proposed various techniques. Traditional rule-based detection methods focus on identifying incomplete or abnormal account details. However, these approaches are limited in accuracy. Modern detection techniques rely on machine learning (ML) and artificial intelligence (AI), where algorithms are trained to differentiate between real and fake accounts using large datasets. Features such as profile information, textual content, activity frequency, and graph/network structure are analyzed to build classifiers. Popular models include Decision Trees, Random Forests, Support Vector Machines, and advanced algorithms like XGBoost and Deep Neural Networks. The importance of fake account detection lies in ensuring user safety, protecting privacy, and maintaining the integrity of social media platforms. By reducing the presence of fake accounts, platforms can prevent the spread of false information, reduce online scams, and promote a healthier digital environment. However, challenges remain, as fake account creators continuously improve their methods to bypass detection systems. Therefore, ongoing research, better datasets, and innovative algorithms are essential to strengthen detection accuracy. This project aims to study and implement effective fake account detection techniques on social media. By analysing user

behaviour, profile details, and activity patterns, the project will demonstrate how intelligent models can classify accounts as genuine or fake. The outcome of this project will contribute to building safer social networks where users can interact with greater trust and confidence. Social media has become widespread in recent times, primarily as a tool for communication, exchange of ideas, and information obtaining.

EXISTING SYSTEM

In the existing system, social media platforms and researchers have been using different techniques to detect and control fake accounts. Traditionally, the methods are mostly rule-based, where platforms define a set of rules and conditions to identify suspicious users. For example, if an account has incomplete profile details, abnormal usernames, or an unusually high number of friend requests in a short period of time, it is flagged as suspicious
These systems are simple and easy to implement, but they are not accurate enough to handle advanced fake accounts that are designed to look like genuine users. Many social media platforms also rely on manual reporting. In this method, genuine users can report suspicious profiles or accounts, which are then reviewed by the platform's moderation team. While this approach helps in removing some fake accounts, it is not scalable because millions of users join social media daily, and it is impossible to manually check all accounts.In recent years, researchers have started to apply basic machine learning models like Decision Trees, Logistic Regression, and Random Forests for fake account detection. These models take input features such as profile data, activity frequency, and interaction history, and then classify accounts as real or fake. Although these approaches are more accurate than rule-based methods, they still face limitations. For example, they require a large and well-labeled dataset for training, and their accuracy decreases when fake accounts change their behaviour over time.

PROPOSED SYSTEM

The proposed system aims to overcome the limitations of existing methods and provide a more reliable, accurate, and intelligent way of detecting fake accounts on social media platforms. Instead of relying only on rule-based checks or manual reporting, this system uses machine learning and data-driven approaches to automatically identify suspicious users with higher precision. The key idea is to extract different types of features from user accounts and use them to train an intelligent model. The proposed approach makes use of advanced machine learning algorithms such as Decision Trees, Random Forests, and particularly XGBoost (Extreme Gradient Boosting). XGBoost is chosen because of its ability to handle large datasets, deal with missing values, reduce overfitting through regularization, and provide high accuracy in classification tasks. The system also emphasizes automation. Once the model is trained, it can automatically scan and classify new accounts in real-time without human intervention. This saves time compared to manual reporting systems and is more scalable to handle millions of accounts simultaneously. Moreover, the system is designed to continuously update itself by retraining on fresh data, so that it can adapt to the evolving tactics of fake account creators. Another important aspect of the proposed system is the inclusion of ethical and privacy considerations. User data is anonymized before processing, ensuring that detection is carried out responsibly without violating user rights. The system aims to minimize false positives (wrongly flagging genuine users) while maximizing true detection rates, so that user trust in the platform is maintained. The proposed system is designed to create a safer and more trustworthy social media environment by accurately detecting fake accounts. By combining machine learning with behavioural and network analysis, it addresses the shortcomings of existing systems and provides an intelligent, efficient, and future-ready solution to one of the most pressing problems in the digital world. Content-based data is also an important part of the system.

LITERATURE SURVERY

The rapid growth of social media platforms has significantly transformed communication, enabling both personal and business interactions, but it has also introduced serious security and privacy challenges. Chen *et al.* [1] analyzed social media usage from a symbolic interactionism perspective, emphasizing how digital communication influences user behavior and information exchange. However, the increasing dependence on social platforms has also led to rising cyber threats. Chen *et al.* [2] highlighted the risks associated with invalid and malicious messages, stressing the need for stronger legal frameworks to combat cybercrime in social applications. Similarly, Jain *et al.* [3] provided a comprehensive review of security and privacy concerns in online social networks, identifying major vulnerabilities such as data leakage, identity theft, and unauthorized access.

A major challenge in social media security is the spread of false information and spam. Guo *et al.* [4] discussed the future of misinformation detection, emphasizing the role of artificial intelligence in identifying fake content. Rao *et al.* [5] reviewed social spam detection techniques and outlined key challenges such as evolving spam patterns and

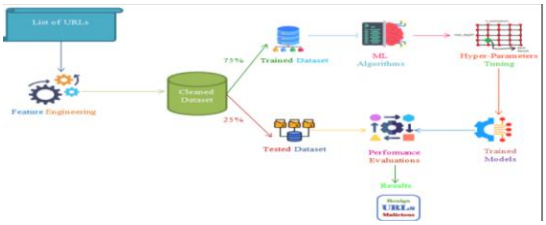
scalability issues. Malware and network-level threats also play a significant role in compromising social media platforms. Caviglione *et al.* [6] presented an overview of modern malware threats and detection trends, while Rabbani *et al.* [7] explored machine learning approaches for detecting malicious behavior in networks, highlighting their effectiveness in identifying anomalies.

Fake profile detection has emerged as a critical research area in social media security. Bharti and Pandey [8] proposed a logistic regression model optimized with particle swarm optimization for detecting fake Twitter accounts, demonstrating improved classification accuracy. Das Gupta *et al.* [9] introduced hybrid feature-based machine learning techniques for detecting phishing websites, which are often linked to fake profiles. Joshi *et al.* [10] and Roy and Chahar [11] provided comprehensive studies on fake profile detection, analyzing various machine learning algorithms and their effectiveness. More recent work by Habib *et al.* [12] explored advanced algorithms for identifying fake profiles, emphasizing the role of deep learning and behavioral analysis.

User privacy and information sharing also remain significant concerns. Hajli and Lin [13] examined how perceived control over information affects user trust and security in social networking sites. Mughaid *et al.* [14] proposed a novel system combining machine learning with face recognition techniques to detect fake accounts, enhancing authentication accuracy. Aljabri *et al.* [15] conducted a comprehensive review of bot detection techniques, highlighting the increasing sophistication of automated social media accounts. Adewole *et al.* [16] further contributed by using clustering and classification methods to detect spam accounts on Twitter.

Beyond technical challenges, social media also plays a role in information warfare and large-scale influence operations. Studies such as [17] and [20] explored how social networks are used as sources of information in conflicts, emphasizing their impact on public opinion and national security. Dudatiev [19] discussed complex information security models and protection mechanisms, while reports like [18] highlighted the widespread adoption of social media platforms, further underlining the importance of robust security measures.
Overall, the literature demonstrates that while machine learning and artificial intelligence have significantly improved the detection of fake profiles, spam, and malicious activities, challenges such as evolving attack strategies, privacy concerns, and large-scale misinformation remain unresolved. Integrating advanced techniques such as behavioral analysis, face recognition, and hybrid models can enhance detection accuracy. These studies collectively provide a strong foundation for developing secure and intelligent systems for social media monitoring and fake account detection.

SYSTEM ARCHITECUTUTE



The system architecture consists of the process for creating machine learning models that can detect dangerous URLs. It outlines the research framework and explains the methods and approaches that were used. The first stage was to prepare the dataset, which came from Kaggle and consisted of a variety of URLs that were classified as either benign or malicious. Any null values in the dataset were meticulously eliminated to guarantee data integrity. Essential characteristics for identifying malicious URLs were then extracted, producing a refined dataset in which every row denoted a distinct URL, identified by its retrieved features and a label designating whether it was malicious or benign.

INPUT AND OUTPUT DESIGN

The input design of the proposed system plays a crucial role in ensuring accurate and reliable detection of fake or malicious accounts. The system is designed to collect large-scale data from various social media platforms, capturing both static and dynamic user information. This includes user profile data, historical activity logs, and interaction patterns. The goal of the input design is to ensure that the dataset reflects real-world scenarios by including diverse behavioural patterns from both genuine and suspicious accounts. A well-designed input system improves the effectiveness of anomaly detection by providing high-quality and relevant data to the model.

The system considers a wide range of input features that represent user behavioral and account characteristics. In addition to these, advanced behavioral features are also extracted, such as posting frequency, time

intervals between consecutive posts, content similarity, and sudden spikes in activity. Features like follower-following ratio, inactive periods followed by bursts of activity, and repeated content posting are particularly useful in identifying bot-like or malicious behavior. The inclusion of both statistical and behavioral features enhances the model's ability to distinguish between normal and anomalous accounts.

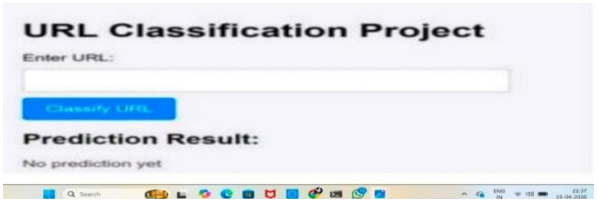
To ensure data quality and consistency, several preprocessing steps are applied to the input data before it is used for model training. These steps include handling missing values through imputation techniques, removing duplicate records, and eliminating noisy or irrelevant data. Feature scaling methods such as normalization and standardization are applied to ensure that all features contribute equally during the training process.

Additionally, categorical variables are transformed into numerical representations using encoding techniques such as label encoding or one-hot encoding. Data validation is also performed to ensure correctness and completeness. This comprehensive preprocessing pipeline ensures that the input data is clean, structured, and suitable for efficient model learning.

RESULTS

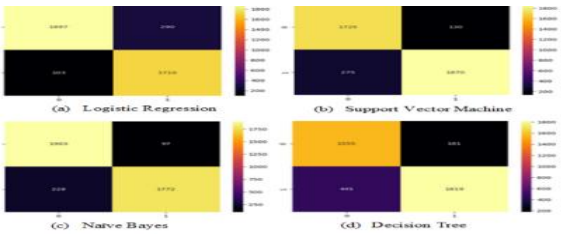
HOME PAGE INTERFACE

LOGIN PAGE INTERFACE

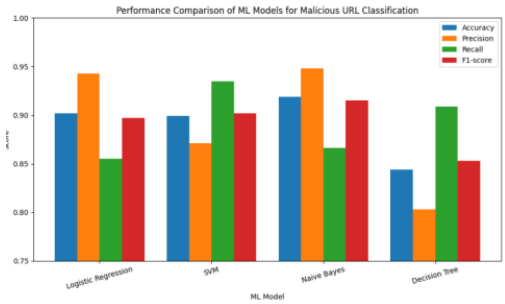


SUPERVISOR INTERFACE

AUTOENCODER DETECTION RESULT INTERFACE



MANUFACTURE INTERFACE



GRAPHICAL ANALYSIS OF DECTION RESULTS

CONCLUSION

The proposed system for detecting malware through autoencoder-driven anomaly analysis provides an intelligent and efficient solution for modern cybersecurity challenges. With the rapid growth of digital systems and increasing cyber threats, traditional malware detection methods are becoming less effective. Signature-based detection techniques fail to identify new and evolving malware attacks, making it necessary to adopt advanced machine learning and deep learning approaches. The proposed system addresses this issue by using an autoencoder-based anomaly detection model that learns normal system behavior and detects suspicious activities. The system begins with

collecting relevant data from system behavior, network activity, and user interactions. After collecting the data, preprocessing techniques are applied to clean and normalize the dataset. Feature extraction plays a crucial role in identifying meaningful attributes that represent system behavior. These features are used to train the autoencoder model. The model learns patterns of normal activity and identifies anomalies based on reconstruction error. This approach enables the detection of unknown malware without requiring labeled datasets. The classification module further enhances detection accuracy by categorizing suspicious activities. The fake account detection module identifies malicious users and abnormal system behavior. Evaluation metrics such as accuracy, precision, recall, and F1-score are used to measure system performance. The testing results demonstrate that the proposed system achieves high detection accuracy and reliability. Another important advantage of the proposed system is its ability to detect zero-day attacks. The system also demonstrates scalability when handling large datasets. This makes it suitable for enterprise-level applications and cloud environments. The proposed system also provides real-time detection capability. Continuous monitoring allows early identification of malware threats. This helps prevent system damage and improves cybersecurity. The system is also flexible and can be integrated with existing security infrastructures. This improves usability and adoption in real-world environments. Despite these advantages, some challenges remain, such as computational complexity and training time. However, these limitations can be addressed through optimization techniques and improved hardware support. Overall, the proposed system provides a reliable and scalable solution for malware detection. The results confirm that autoencoder-driven anomaly detection is an effective approach for improving cybersecurity. The system enhances threat detection, reduces cyber risks, and improves system security in modern digital environments.

FUTURE ENCHANCEMENT

Future enhancements to the proposed malware detection system includes several enhancements to improve performance, scalability, and real-world deployment. One of the major future improvements is integrating advanced deep learning models such as variational autoencoders and hybrid deep learning architectures. These advanced models can further improve detection accuracy and reduce false positives. Future work may also focus on combining autoencoder models with reinforcement learning techniques for adaptive malware detection. This improves transparency and trust in the system. Future research may also focus on improving visualization dashboards to display anomaly detection results more effectively. Another future enhancement involves reducing computational complexity. Optimized architectures and lightweight models can be developed for deployment in resource-constrained environments. This includes mobile devices, embedded systems, and edge computing environments. Such improvements will increase system applicability across various platforms. The proposed system can also be extended to detect different types of cyber threats such as phishing attacks, ransomware, and insider threats. Integrating multi-layer security frameworks will further improve detection performance. Future work may also include implementing automated response mechanisms that take immediate action when malware is detected. This reduces manual intervention and improves system security. Additionally, future research may focus on improving dataset quality by collecting large-scale real-world datasets. Continuous learning mechanisms can be implemented to retrain the model periodically. This ensures that the system adapts to evolving malware patterns. The system can also be integrated with cloud-based security solutions for centralized monitoring. Overall, the future scope of the proposed system is vast and promising. Continuous advancements in artificial intelligence and cybersecurity will further enhance malware detection capabilities. The proposed system can be extended to multiple domains and environments. Future improvements will make the system more robust, scalable, and efficient. These enhancements will contribute to stronger cybersecurity solutions and improved protection against evolving cyber threats.

REFERENCE

[1] R. R. Chen, R. M. Davison, and C. X. Ou, “A symbolic interactionism perspective of using social media for personal and business communication,” International Journal of Information Management, vol. 51, p. 102022, Apr. 2020.

[2] H. Chen, Y. Yang, and Y. Wu, “Invalid message risks and analysis of laws to restrict cyber crime in social applications,” in Proc. 11th Int. Conf. Networks, Communication and Computing, New York, USA: ACM, Dec. 2022, pp. 341–347.

[3] A. K. Jain, S. R. Sahoo, and J. Kaubiyal, “Online social networks security and privacy: Comprehensive review and analysis,” Complex Intelligent Systems, vol. 7, no. 5, pp. 2157–2177, 2021.

[4] B. Guo, Y. Ding, L. Yao, Y. Liang, and Z. Yu, “The future of false information detection on social

media," ACM Computing Surveys, vol. 53, no. 4, pp. 1–36, Jul. 2021.

[5] S. Rao, A. K. Verma, and T. Bhatia, "A review on social spam detection: Challenges, open issues, and future directions," Expert Systems with Applications, vol. 186, p. 115742, Dec. 2021.

[6] L. Caviglione et al., "Tight arms race: Overview of current malware threats and trends in their detection," IEEE Access, vol. 9, pp. 5371–5396, 2021.

[7] M. Rabbani et al., "A review on machine learning approaches for network malicious behavior detection in emerging technologies," Entropy, vol. 23, no. 5, p. 529, Apr. 2021.

[8] K. K. Bharti and S. Pandey, "Fake account detection in Twitter using logistic regression with particle swarm optimization," Soft Computing, vol. 25, no. 16, pp. 11333–11345, Aug. 2021.

[9] S. Das Guptta et al., "Modeling hybrid feature-based phishing websites detection using machine learning techniques," Annals of Data Science, vol. 11, no. 1, pp. 217–242, Feb. 2024.

[10] U. D. Joshi et al., "Fake social media profile detection," in Machine Learning Algorithms and Applications, Wiley, 2021, pp. 193–209.

[11] 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.

[12] A. K. M. R. R. Habib et al., "Techniques to detect fake profiles on social media using new age algorithms: A survey," in Proc. IEEE CCWC, 2024, pp. 0329–0335.

[13] N. Hajli and X. Lin, "Exploring the security of information sharing on social networking sites: The role of perceived control of information," Journal of Business Ethics, vol. 133, no. 1, pp. 111–123, 2021.

[14] A. Mughaid et al., "A novel machine learning and face recognition technique for fake accounts detection system," Multimedia Tools and Applications, vol. 82, no. 17, pp. 26353–26378, Jul. 2023.

[15] M. Aljabri et al., "Machine learning-based social media bot detection: A comprehensive literature review," Social Network Analysis and Mining, vol. 13, no. 1, p. 20, Jan. 2023.

[16] K. S. Adewole et al., "Twitter spam account detection based on clustering and classification methods," Journal of Supercomputing, vol. 76, no. 7, pp. 4802–4837, Jul. 2020.

[17] "Information warfare: The role of social media in conflict," UNT Digital Library.

[18] "The 15 biggest social media sites and apps," Dreamgrow, 2022.

[19] A. V. Dudatiev, Complex Information Security of STS: Models of Influence and Protection. VNTU, 2022.

[20] O. P. Voitovych and V. O. Holovenko, "Research of social networks as a source of information in warfare," in Engineering of XXI Century, pp. 111–119, 2020.