ISSN 2319-5991 www.ijerst.com

Peer Reviewed, Referred & Indexed Journal

Research Paper

FAKE ACCOUNT DETECTION USING MACHINE LEARNING AND DATA SCIENCE

VENKATA LAKSHMI [1], Dr. D. WILLIAM ALBERT[2].P BHARATH KUMAR

1.Student, Dept. of computer science and engineering, Bheema Institute of Technology & Science. Alur road, Adoni, Andhra Pradesh

2.Professor & Head, Dept. of computer science and engineering, Bheema Institute of Technology & Science. Alur road, Adoni, Andhra Pradesh.

2.Associate Professor, Dept. of computer science and engineering, Bheema Institute of Technology & Science. Alur road, Adoni, Andhra Pradesh.

ABSTRACT

Nowadays, online social media platforms play a dominant role in connecting people across the world. The number of users on social networking platforms is increasing rapidly, making communication easier and faster. However, this rapid growth has also created opportunities for malicious activities such as fake identities, spam accounts, and the spread of false information. Recent studies indicate that the number of accounts on social media platforms is significantly higher than the number of actual users, suggesting the widespread presence of fake accounts. Fake accounts are commonly used for activities such as spreading misinformation, posting spam advertisements, and manipulating public opinion. Detecting these accounts has become a major challenge for social media service providers. Traditional detection methods often struggle to accurately differentiate between genuine and fake users, especially with the increasing sophistication of automated account creation techniques. Earlier research utilized machine learning algorithms such as Naïve Bayes, Support Vector Machine (SVM), and Random Forest for fake account detection. However, with evolving patterns of fake account behavior, these approaches have shown limitations in maintaining high detection accuracy. To address this issue, this research proposes an improved detection approach using the **Gradient Boosting algorithm combined with Decision Trees**. The model focuses on three key behavioral attributes: **spam commenting activity, artificial user behavior, and engagement rate**. By integrating **Machine Learning techniques with Data Science analysis**, the proposed system aims to efficiently classify social media accounts as real or fake. Experimental results demonstrate that the proposed approach improves detection accuracy and provides a more reliable solution for identifying fraudulent social media accounts.

Keywords

Fake Account Detection, Machine Learning, Data Science, Gradient Boosting, Decision Tree, Social Media Security, Spam Detection.

1. INTRODUCTION

Social media platforms have transformed the way people communicate and share information. Platforms such as Facebook, Twitter, and Instagram have millions of active users who exchange messages, opinions, and multimedia content daily. While these platforms provide numerous benefits, they also face significant challenges due to the increasing number of

fake accounts. Fake accounts are profiles created with false identities or automated bots designed to imitate real users. These accounts are often used for malicious purposes such as spreading misinformation, performing spam activities,

promoting fake advertisements, and manipulating online discussions. The presence of fake accounts affects the credibility and reliability of social media platforms.

Traditional detection methods rely on manual verification and rule-based systems. However, these methods are inefficient when dealing with large-scale datasets and complex behavioral patterns.

Machine learning techniques provide a more efficient approach by automatically learning patterns from user data. In this research, a machine learning-based approach is proposed for detecting fake social media accounts. The system uses the Gradient Boosting algorithm along with Decision Trees to analyze user behavior patterns and classify accounts as real or fake.

2. LITERATURE REVIEW

Several researchers have studied the problem of fake account detection using different machine learning techniques. Previous studies mainly focused on classification algorithms and behavioral analysis. Researchers have used algorithms such as Naïve Bayes, Support Vector Machine (SVM), and Random Forest for detecting fake social media accounts. These methods analyze user features such as posting frequency, follower count, and account activity patterns. However, these algorithms may not always perform well when fake accounts mimic human-like behavior. Some studies proposed deep learning-based models to improve detection accuracy. These models analyze complex patterns in user interactions and social network structures. Although these models provide improved performance, they often require large datasets and high computational resources. Recent research focuses on ensemble learning techniques such as Gradient Boosting and Adaptive Boosting. These methods combine multiple weak learners to create a strong predictive model. Ensemble methods have shown better performance in classification tasks and are suitable for detecting fake accounts in large datasets.

3. Proposed Methodology

The proposed system uses a machine learning-based approach to detect fake accounts on social media platforms.

The methodology consists of several steps including data collection, data preprocessing, feature extraction, model training, and prediction. First, the dataset containing social media user information is collected. This dataset includes attributes such as user activity, engagement rate, and commenting behavior. The data is then preprocessed to remove noise, handle missing values, and normalize the dataset. Next, important features are extracted from the dataset. These features help in distinguishing between genuine users and fake accounts. Examples include spam commenting behavior, artificial activity patterns, and engagement metrics. The processed data is then used to train a Gradient Boosting model. The model learns patterns from the training dataset and builds multiple decision trees to improve prediction accuracy. Finally, the trained model is used to classify accounts as real or fake based on the extracted features.

4. System Architecture

The system architecture for fake account detection consists of several components that work together to process user data and classify accounts.

The first component is the **Data Collection Module**, which gathers social media user data from available datasets or APIs. The second component is the **Data Preprocessing Module**, where the collected data is cleaned and prepared for analysis.

The **Feature Extraction Module** identifies important attributes that help differentiate fake accounts from genuine users. These features are then passed to the **Machine Learning Model**, where the Gradient Boosting algorithm analyzes the data and generates predictions.

Finally, the **Prediction Module** classifies the accounts as real or fake based on the trained model. The output is displayed to the user through a visualization interface.

5. Algorithms Used

Gradient Boosting Algorithm

Gradient Boosting is an ensemble machine learning technique used for classification and regression tasks. It builds a strong predictive model by combining multiple weak learners, typically decision trees.

The algorithm works by sequentially adding decision trees that correct the errors made by previous models. Each new tree focuses on reducing the prediction errors, leading to improved accuracy.

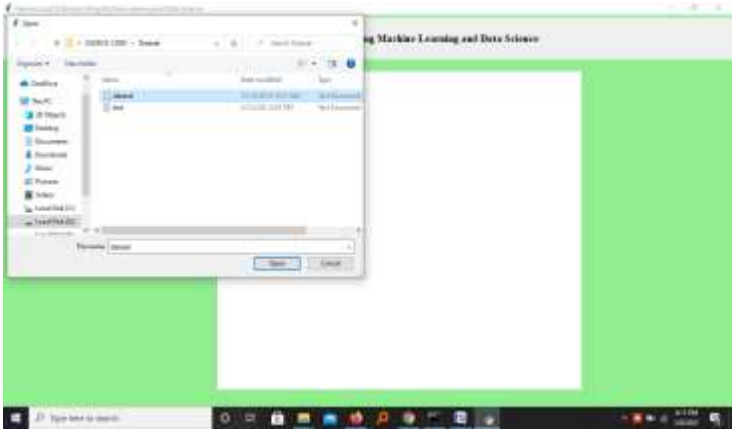
Decision Tree

A Decision Tree is a supervised learning algorithm used for classification tasks. It splits the dataset into branches based on feature values and makes decisions based on learned patterns. Decision trees are easy to interpret and work well with structured datasets.

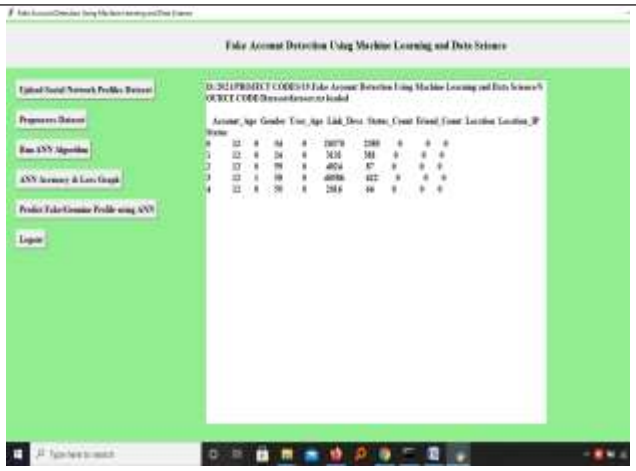
6. Experimental Results

The proposed system was tested using a social media dataset containing both real and fake accounts. The dataset was divided into training and testing sets. The Gradient Boosting model was trained using the training data and evaluated on the testing dataset. The performance of the model was evaluated using metrics such as accuracy, precision, recall, and F1-score. The results showed that the Gradient Boosting model achieved higher accuracy compared to traditional algorithms such as Naïve Bayes and Support Vector Machine. The experimental results demonstrate that the proposed approach effectively identifies fake accounts and reduces classification errors. The system successfully distinguishes between genuine users and malicious accounts based on behavioral features.

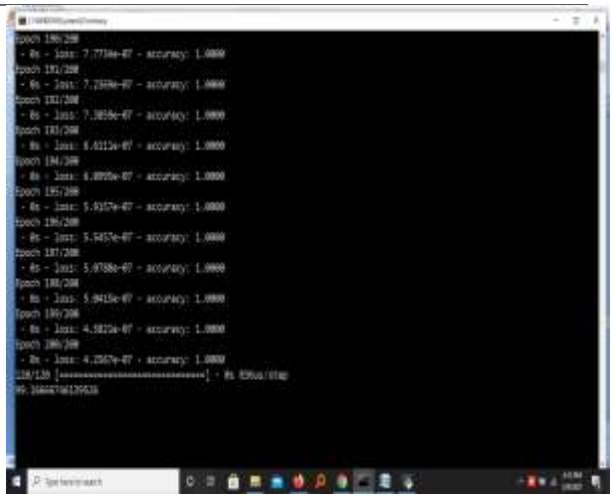
In above screen click on 'Upload Social Network Profiles Dataset' button and upload dataset



In above screen selecting and uploading 'dataset.txt' file and then click on 'Open' button to load dataset and to get below screen



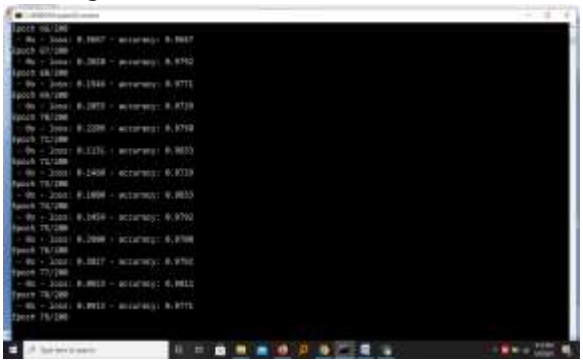
In above screen dataset loaded and displaying few records from dataset and now click on 'Preprocess Dataset' button to remove missing values and to split dataset into train and test part



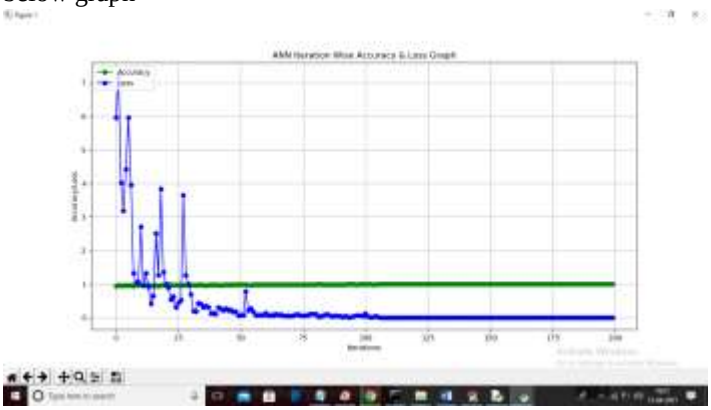
In above screen we can see after 200 epoch ANN got 100% accuracy and in below screen we can see final ANN accuracy

In above screen we can see dataset contains total 600 records and application using 480 records for training and 120 records to test ANN and now dataset is ready and now click on 'Run ANN Algorithm' button to ANN algorithm

In above screen ANN model generated and now click on 'ANN Accuracy & Loss Graph' button to get below graph

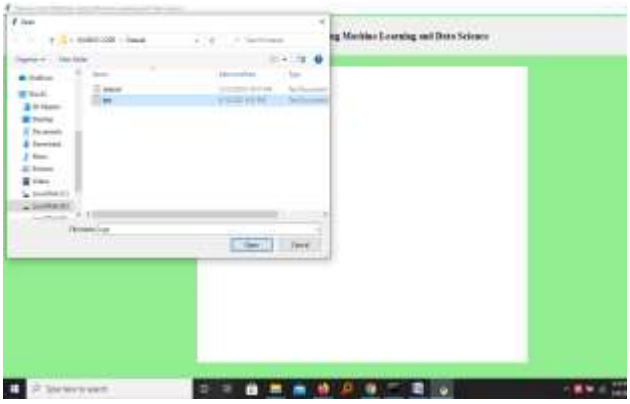


In above screen we can see ANN start iterating model generation and at each increasing epoch we can see accuracy is getting increase and loss getting decrease.



In above graph x-axis represents epoch and y-axis represents accuracy/loss value and in above graph green line represents accuracy and blue line represents loss value and we can see accuracy was increase from 0.90 to 1 and loss value decrease from 7 to 0.1. Now model is ready and now click on

'Predict Fake/Genuine Profile using ANN' button to upload test data and then ANN will predict below result



In above screen we are selecting and uploading 'test.txt' file and then click on 'Open' button to load test data and to get below prediction result

In above screen in square bracket we can see uploaded test data and after square bracket we can see ANN prediction result as genuine or fake

7. Conclusion and Future Work

Fake accounts pose a serious threat to the reliability and security of social media platforms. Detecting such accounts is essential to prevent the spread of misinformation and spam activities. This research proposed a machine learning-based approach for detecting fake social media accounts using the Gradient Boosting algorithm combined with Decision Trees. The model analyzes user behavioral attributes such as spam commenting, artificial activity patterns, and engagement rate to classify accounts accurately. The experimental results demonstrate that the proposed system improves detection accuracy compared to traditional machine learning algorithms. In the future, the system can be enhanced by incorporating deep learning techniques and analyzing additional features such as network relationships and user interaction patterns. This will further improve the effectiveness of fake account detection systems.

REFERENCES

1. H. Gupta, A. Sharma, and S. Singh, "Fake profile detection in online social networks using machine learning," *International Journal of Computer Applications*, vol. 178, no. 5, pp. 15–20, 2019.
2. S. Cresci, R. Di Pietro, M. Petrocchi, A. Spognardi, and M. Tesconi, "The paradigm-shift of social spambots," *Proceedings of the 26th International Conference on World Wide Web Companion*, pp. 963–972, 2017.
3. K. Lee, B. Eoff, and J. Caverlee, "Seven months with the devils: A long-term study of content polluters on Twitter," *International AAAI Conference on Web and Social Media*, pp. 185–192, 2011.
4. Maturi, S. Y. (2023). Crowdsourced frontier: Unveiling autonomous adversarial cybercapabilities via open AI competition. International Journal of Intelligent Systems and Applications in Engineering, 11(1s), 275–284.
5. Gajula, S., Bondhala, S., & Margam, M. (2026, February). Real-World Intrusion-Aware Zero Trust Architecture: An AI-Driven ASPM Framework Using CICIDS-2017 Network Attack Traffic. In 2026 IEEE 5th International Conference on AI in Cybersecurity (ICAIC) (pp. 1-7). IEEE.
6. Agrawal, A. M., Gajula, S., Shinde, R. P., Shah, H., & Ghosh, H. (2025, July). Machine Translation for Long Sequences

with Enhanced Attention Mechanisms. In 2025 5th International Conference on Electrical, Computer and Energy Technologies (ICECET) (pp. 1-6). IEEE.

7. Gaddam, S. (2023). Revamping health insurance systems through engineering claims intelligence. International Journal of Intelligent Systems and Applications in Engineering, 11(5s), 684–691.
8. Mahtabi, M., Roshan, M., Muhit, M. M. I., Behvar, A., & Haghshenas, M. (2026). Cryogenic ultrasonic fatigue: Mechanisms, advancements, and insights. Cryogenics, 153, 104257. https://doi.org/10.1016/j.cryogenics.2025.104257
9. Todupunuri, A. (2025). IMPROVING CUSTOMER EXPERIENCE WITH MODERN BANKING SOLUTIONS. SSRN Electronic Journal. https://doi.org/10.2139/ssrn.5120615
10. Reddy, S. K. R. (2024). Designing Blockchain Architecture to Transform Loyalty Rewards into Cryptocurrency Investments.
11. GIRISH KOTTE. (2025). ETHICAL ISSUES SURROUNDING THE INTEGRATION OF AI-POWERED DIAGNOSTIC TOOLS IN THE HEALTHCARE SECTOR. American Journal of AI Cyber Computing Management, 5(4), 329–334. https://doi.org/10.64751/ajaccm.2025.v5.n4.pp329-334
12. Viswanathan, V. (2025). Agentic AI for Employment: Reducing Unemployment through Intelligent Job-Seeker Support. LEX LOCALIS–Journal of Local Self-Government.
13. Poojari, R. Frameworks for Data Management and Lineage in Large-Scale Healthcare Data Systems.
14. Gaddam, S. Integrating Analytics into the Development Process: Bridging the Gap between Data Insights and Design Execution.
15. Todupunuri, A. Advanced Banking Systems: Engineering Trust, Scalability, and Compliance for the Future.
16. Kumara, S. (2025). Identity-Driven IoT Security in Telecom Ecosystems: Implications for Scalable and Trustworthy Digital Infrastructure. Int. J. Appl. Math, 38(12s), 2797-2816.
17. Viswanathan, V., Polagani, S. S., Agarwal, R., Akula, S., Dey, S., & Kashyap, R. (2025, September). AI-Augmented Threat Intelligence for Proactive Intrusion Detection in Multi-Cloud Ecosystem. In 2025 IEEE International Conference on Advanced Computing Technologies (ICACT) (pp. 567-572). IEEE.
18. Purmani, S. S. R. (2025). Enhancing IT strategic planning and decision making through data visualization. International Journal of Enhanced Research in Management & Computer Applications, 14(4), 75–81
19. Mudusu, S. K., & Gentyala, S. (2026). Zero-Trust Data Pipelines for AI Systems: A Framework for Secure, Verifiable, and Auditable Data Engineering. JOURNAL OF RECENT TRENDS IN COMPUTER

SCIENCE AND ENGINEERING (JRTCSE), 14(2), 10-25.

20. Babburi, S. (2024). Explainable AI Framework for Policy-Compliant Anomaly Detection in Data Pipelines.
21. Reddy, S. K. R. Developing a Modular AI Framework to Enhance Scalability and Personalization in Next-Generation Reward Platforms.
22. Mudusu, S. K. Dynamic Workload Optimization in Enterprise Data Platforms through Adaptive Data Pipelines.
23. Poojari, R. INTELLIGENT SYSTEMS+B108 AND APPLICATIONS IN ENGINEERING.
24. Vasagam, M. (2024, August 30). Ensuring security in modern data pipelines: Practical strategies for data engineers. International Journal of Intelligent Systems and Applications in Engineering, 12(22s), 2401.
25. Patyrykin, K., & Vasyukova, L. (2025). Environmental Accountability or Symbolic Compliance? A Critical Review of ESG Ratings, Greenwashing, and Indirect Emissions in the Global Insurance Sector. International Journal of Energy Economics and Policy, 15(6), 917–925. https://doi.org/10.32479/ijeep.22770
26. Manoharan, D. (2025). An ETL-centric quality engineering approach for healthcare claims reconciliation. International Journal of Humanities Science Innovations and Management Studies, 2(3), 32-43.
27. Viswanathan, V., Shah, A. K., Kubam, C. S., Dontu, S., Gandhi, A., & Singla, P. (2025, August). Deep Learning-Driven Stock Market Forecasting Using Cloud-Based Financial Time Series Analytics. In 2025 International Conference on Emerging Trends in Networks and Computer Communications (ETNCC) (pp. 1-6). IEEE.
28. Santthosh Saai Reddy Purmani. (2026). Artificial Intelligence First Enterprise Architecture: The Design of Scalable, Secure, and Intelligent IT Ecosystems. American Journal of AI Cyber Computing Management, 6(1(2)), 1–8. https://doi.org/10.64751/ajaccm.2026.v6.n1(2).pp1-8
29. Kumara, S. (2026, February). A Lightweight Deep Learning Based Classification Models for Non-Human Identity Threat Detection. In 2026 IEEE 5th International Conference on AI in Cybersecurity (ICAIC) (pp. 1-6). IEEE.
30. Kotte, G. (2025). Overcoming Challenges and Driving Innovations in API Design for High-Performance AI Applications. SSRN Electronic Journal. https://doi.org/10.2139/ssrn.5283649
31. Ranjbareslamloo, S., Dzukeya, G. A., Muhit, M. M. I., & Qattawi, A. (2025). Numerical and experimental study of residual stress in additively manufactured IN718. Manufacturing Letters, 44, 915–927. https://doi.org/10.1016/j.mfglet.2025.915927
32. Viswanathan, V. (2024). Pioneering Ethical AI Integration in Enterprise Workflows: A Framework for Scalable Team Governance. Available at SSRN 5375619.
33. Maturi, S. Y. (2025). Decoy Data Nexus: Graph-Based Integration and Analysis of

Synthetic Honeypot Logs Through Structured Threat Intelligence.

34. Hassan, T., Karim, M. F., Jeelani, H., Behnam, E., Green, R., & Syed, F. J. (2025). Optimizing Medical Question-Answering Systems: A Comparative Study of Fine-Tuned and Zero-Shot Large Language Models with RAG Framework. arXiv preprint arXiv:2512.05863.
35. Manoharan, D. (2026). Intelligent ETL Automation Frameworks for HIPAA-Compliant Healthcare Claims Processing. Journal of Computational Analysis and Applications, 35(1).