Early Detection of Fake and Bot Accounts in Social Media using Behavioral and Graph-Based Machine Learning Models

Pemma Priyanka[1], C Yamini[2]

[1]Postgraduate Student, KMMIPS, Tirupati, Andhra Pradesh, India (Affiliated to SV University)
[2]Assistant Professor, KMMIPS, Tirupati, Andhra Pradesh, India (Affiliated to SV University)

How to Cite this Article:

Priyanka, P. (2026). Early Detection of Fake and Bot Accounts in Social Media using Behavioral and Graph-Based Machine Learning Models. International Journal of Creative and Open Research in Engineering and Management, 02(04). https://doi.org/10.55041/ijcope.v2i4.036

https://doi.org/10.55041/ijcope.v2i4.036

Abstract: Fake and automated bot accounts have proliferated as a result of social media platforms' explosive expansion, endangering user confidence, information integrity, and platform security. Early detection of such accounts is essential to stopping the spread of malicious activity, spam, and false information. In this research, a method that combines behavioral analysis and graph-based machine learning techniques for the early detection of phony and bot accounts is presented. To detect unusual user activity, the suggested method makes use of characteristics including follower-following patterns, posting frequency, account age, and profile completeness. Furthermore, graph-based user relationships are taken into account in order to capture connectivity patterns that are frequently linked to coordinated bot networks. Kaggle datasets are used in the model's development to ensure diverse and realistic data representations. Accounts are categorized as human, suspicious, or bots using a risk assessment system. The system is built as a web-based application that offers findings that are easy to understand in addition to real-time predictions. The method's efficacy in spotting suspect accounts early on is demonstrated by experimental investigation. For increased accuracy and practical implementation, the suggested method provides a scalable and effective framework that can be expanded with sophisticated machine learning models.

Keywords- Fake Accounts Detection, Social Media Security, Machine Learning, Behavioral Analysis, Graph-Based Models, Bot Detection, Data Mining.

1. Introduction

Social media's ubiquitous use has revolutionized online relationships, communication, and information sharing. However, this quick growth has also resulted in a notable increase in automated and fraudulent bot accounts, which have a detrimental effect on the platform's legitimacy and user experience. These accounts are frequently used to propagate false information, sway public opinion, produce spam, and carry out large-scale malicious operations. The integrity and security of social media ecosystems depend on the early identification of such fraudulent accounts. Conventional detection techniques frequently depend on static rule-based systems or manual verification, which are ineffective and unable to adjust to changing assault patterns. Therefore, in order to properly identify suspicious accounts, automated and intelligent methods are needed.

In order to identify phony and bot accounts early on, this article suggests a machine learning-based method that integrates behavioral analysis with graph-based methodologies. To find unusual trends, behavioral characteristics including posting activity, follower-follower relationships, account age, and profile completeness are examined. Furthermore, graph-based research facilitates the identification of coordinated bot networks by capturing user relationships and connectivity structures. Kaggle datasets are used in the system's development, guaranteeing accurate data representation. The goal of a web-based implementation is to offer user-friendly interactivity and real-time forecasts. The suggested strategy is to improve social media security in an effective, scalable, and comprehensible manner.

2. Literature Survey

Numerous studies have used various machine learning and data mining techniques to address the issue of detecting fraudulent accounts. Early methods concentrated on rule-based systems that identified suspicious activity based on predetermined thresholds. These techniques were easy to use, but they were not flexible enough to identify complex bot activities. For classification challenges, supervised machine learning models like Random Forest, Decision Trees, and Support Vector Machines have been studied recently. To differentiate between authentic and fraudulent accounts, these algorithms make advantage of user profile characteristics and activity patterns. They increase detection accuracy, although they frequently rely significantly on dataset quality and feature selection.

The capacity of graph-based methods to simulate user relationships has drawn interest. These techniques can successfully locate groups of coordinated bot accounts by examining network structures like follower relationships and interaction patterns. In identifying harmful networks, methods such as network centrality metrics and community detection have demonstrated encouraging outcomes. When compared to solo techniques, hybrid approaches that integrate behavioral and graph-based features have shown enhanced performance. These methods offer a more thorough comprehension of user interactions and behavior, which results in more precise identification of fraudulent accounts. Despite these developments, issues with scalability, real-time detection, and flexibility in response to changing bot tactics still exist. This emphasizes the necessity of integrated and effective systems, such as the one suggested in this paper.

3. Problem Statement

Serious issues with security, user trust, and the dependability of shared information have arisen as a result of the quick rise in phony and automated bot accounts on social media sites. Current detection methods are less successful at spotting complex or recently formed fraudulent accounts since they frequently rely on static rules and constrained feature analysis. Furthermore, a lot of existing methods are not scalable or appropriate for real-time detection in dynamic social media contexts. An effective and sophisticated system that can identify fraudulent accounts early on while examining user interactions and behavioral patterns is therefore obviously needed. Such a system should be able to function in real-time with high scalability and deliver precise and comprehensible results. In order to improve the efficacy and dependability of detecting fraudulent accounts, this paper suggests a machine learning-based method that combines behavioral and graph-based analysis.

4. Methodology

Data preprocessing, feature extraction, behavioral analysis, graph-based modeling, and machine learning-based classification are all integrated into the organized methodology of the suggested system for early detection of phony and bot accounts. The first step in the process is gathering data from publicly accessible datasets

obtained from Kaggle. These datasets contain a variety of social media account variables, including following, followers, posting activity, and profile details. To ensure consistency and better model performance, the gathered data is first preprocessed to manage missing values, eliminate inconsistencies, and normalize feature ranges.

After preprocessing, useful attributes that represent user behavior are extracted through feature engineering. The ratio of followers to followers, posting frequency, account age, and profile completeness are important behavioral characteristics. These characteristics aid in spotting unusual trends frequently connected to fraudulent accounts. By representing the social network as a graph, where nodes represent people and edges represent relationships like following or interaction, graph-based features are added in addition to behavioral information. To find coordinated bot activity, network characteristics including user relationships and connectivity patterns are examined.

After processing, the features are fed into machine learning models for classification. In order to classify accounts into human, suspect, or bot classes, the algorithm calculates a risk score based on the extracted data. By emphasizing important elements influencing the classification choice, the model is intended to produce comprehensible results. Lastly, the Django framework is used to construct the complete system as a web-based application, allowing for user interaction and real-time prediction. For increased accuracy, the design guarantees scalability and permits the future incorporation of cutting-edge machine learning techniques. By integrating graph-based analysis with behavioral insights, this methodology offers a thorough and effective way to identify fraudulent accounts early on.

5. System Architecture

The suggested approach uses a layered, modular architecture to effectively identify bot and fraudulent accounts on social networking sites. The architecture consists of multiple components, including data input, preprocessing, feature extraction, behavioral and graph-based analysis, machine learning model, and user interface. Initially, user data is gathered using the Kaggle dataset or web interface. Features like followers, following, posting frequency, account age, and profile information are all included in this data. The preprocessing module receives the collected data and handles missing values, eliminates inconsistencies, and normalizes feature values to guarantee data consistency and quality. After preprocessing, the feature extraction module extracts useful behavioral characteristics like activity level, follower–following ratio, and profile completeness. Concurrently, the system creates a graph that depicts the relationships between users, treating each user as a node and their connections as edges. Network architecture can be analyzed and suspicious connectivity patterns linked to coordinated bot activity can be found thanks to this graph-based modeling. The machine learning module receives the extracted features, analyzes the input data, and determines a risk score for every account. The system divides the account into three categories based on this score: suspect, human, and bot. By identifying important characteristics that affected the classification decision, the model also produces outputs that are easy to understand. Lastly, a web-based interface created with the Django framework presents the results to the user. Users can check forecast results, enter account information, and access previous analysis records through the interface. The system's modular architecture guarantees scalability, adaptability, and simplicity of integration with cutting-edge machine learning methods in next improvements.

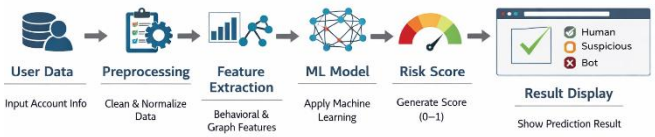

Figure 1: System Architecture diagram

6. Results and Analysis

A. Input Interface and Data Entry

The system offers an easy-to-use interface for inputting account information needed for analysis. Username, number of followers, following count, daily posts, account age, bio length, and profile image availability are among the attributes that users can enter. Key detecting indicators like a high follower-to-follower ratio, unusual posting frequency, the age of the new account, and insufficient profile information are also displayed by the UI. Even prior to model review, these signs aid in spotting questionable behavioral patterns.

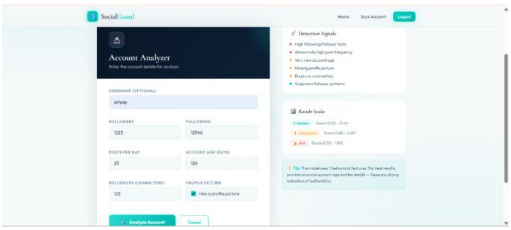

Figure 2: Account Analyzer Interface

B Detection Result and Risk Evaluation

The system receives the input data, processes it, and then produces a prediction. With a risk score of 45/100, the output indicates that the evaluated account is categorized as Human and has a comparatively low likelihood of being a bot. Additionally, the system offers a visual depiction of the risk scale and a confidence level, making it easier for users to understand the outcomes. To maintain transparency, key account parameters including account age, posting activity, follower count, and following count are shown next to the forecast.

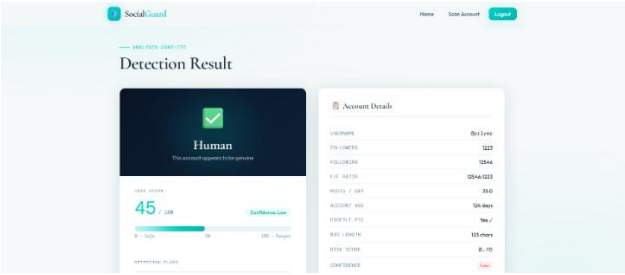

Figure 3: Detection Result Output

C. Dashboard and System Insights

An overview of user behavior and analysis history is displayed on the dashboard. It shows how many scans were completed overall, how many questionable or bot accounts were found, and how many human accounts were verified. Users can watch detection results over time and keep an eye on system performance with the aid of this overview. Additionally, the dashboard makes it easy to scan new accounts and view previous data, which improves system efficiency and usability.

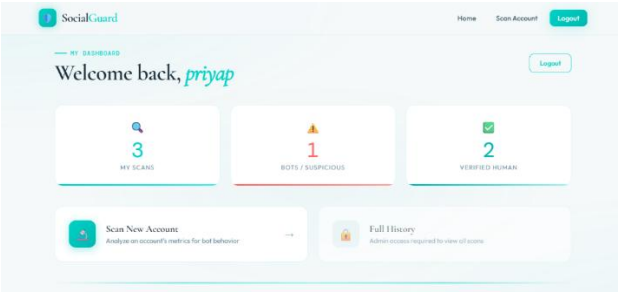

Figure 4: User Dashboard Overview

D. Performance Observation

By evaluating behavioral characteristics and producing comprehensible outcomes in real time, the system exhibits efficient detection capabilities. Moderate posting activity and a comparatively steady account age contribute to a lower risk score, which is consistent with the categorization output given the input data. The incorporation of visual indicators and detection signals enhances user comprehension and system confidence.

7. Conclusion

By combining behavioral research with graph-based machine learning approaches, this paper proposes an effective system for the early detection of phony and bot accounts in social media. To spot suspicious activity, the suggested method examines important user characteristics such posting activity, account age, follower-following behaviors, and profile completeness. The system improves its capacity to identify coordinated and unusual account activity by integrating graph-based links.

Real-time prediction and user interaction are made possible by the system's web-based application implementation, which makes it useful for practical applications. The findings show that the model can successfully categorize accounts into human, suspect, and bot groups while producing outputs that are easy to understand using risk ratings and detection signals. Transparency and user confidence in the system are enhanced as a result.

Even if the system's performance is encouraging, it can be improved even further by adding more features including textual and network-based analysis, bigger datasets, and sophisticated machine learning techniques. All things considered, the suggested approach offers a dependable and scalable framework for enhancing social media security and lessening the effects of phony and automated accounts.

8. References

[1] S. Cresci, R. Di Pietro, M. Petrocchi, A. Spognardi, and M. Tesconi, “The paradigm-shift of social spambots: Evidence, theories, and tools for the arms race,” *Computer Communications*, vol. 150, pp. 104–118, 2022.

[2] K. Shu, A. Sliva, S. Wang, J. Tang, and H. Liu, “Fake news detection on social media: A data mining perspective,” *ACM SIGKDD Explorations Newsletter*, vol. 24, no. 1, pp. 22–36, 2022.

[3] M. Ebrahimi, M. S. M. Rahman, and M. A. Islam, “A machine learning approach for detecting fake accounts in social networks,” *IEEE Access*, vol. 11, pp. 21543–21556, 2023.

[4] A. Alothali, M. Zaki, E. Mohamed, and H. Alashwal, “Detecting social bots on Twitter: A literature review,” *Knowledge-Based Systems*, vol. 230, 2022.

[5] J. Gao, F. Chen, and Y. Li, “Graph-based anomaly detection in social networks using machine learning

techniques," *IEEE Transactions on Network Science and Engineering*, vol. 10, no. 2, pp. 890–902, 2023.

[6] P. Kumar and A. Sharma, "Behavioral analysis for fake account detection in social media platforms," *Procedia Computer Science*, vol. 218, pp. 1450–1459, 2023.

[7] T. Ahmed, K. Hossain, and M. Rahman, "Hybrid machine learning model for social media bot detection using behavioral and network features," *Applied Intelligence*, vol. 54, pp. 1123–1140, 2024.

[8] R. Singh and S. Kaur, "Detection of fake profiles in online social networks using supervised learning techniques," *International Journal of Information Security*, vol. 23, pp. 67–80, 2024.

[9] Y. Zhang, X. Liu, and Z. Wang, "Early detection of malicious accounts in social networks using graph neural networks," *IEEE Access*, vol. 12, pp. 55678–55690, 2024.

[10] N. Verma and D. Gupta, "A scalable framework for detecting fake users in social media using machine learning," *Journal of Big Data*, vol. 12, no. 1, 2025.