ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/IJISRT24NOV108

Fake Social Media Profile Detection and Reporting Using Blockchain Technology

Students: [1] Shreyash Deshmukh; [2] Gaurav Sati; [3] Simran Kalaska ; [4] Mansha Gupta.

Guide: [5] Saurabh Vyas (Assistant Prof.)
Department of Computer Science & Engineering (Cyber security),
GHRCEM, Nagpur

Abstract:- The proliferation of fake social media profiles poses a significant threat to user security, trust, and the integrity of online platforms. This project proposes a novel system that leverages machine learning and blockchain technology to combat this growing problem. The system aims to accurately identify fake profiles by analyzing user behavior, content, and relevant metrics. Blockchain technology is employed to ensure data integrity, transparency, and immutability in profile verification and reporting. A secure and transparent reporting mechanism is developed using blockchain to log and track reports of fake profiles. The project aims to enhance user trust by reducing the prevalence of fake profiles and creating a more trustworthy online environment. The system also facilitates secure data sharing between platforms to enable collaboration in combating fake profiles while protecting user privacy. The expected results include an accurate detection system, an immutable reporting platform, and enhanced trust and security on social media platforms. This project represents a significant step forward in addressing the challenges posed by fake social media profiles and creating a more secure and trustworthy online ecosystem.

I. INTRODUCTION

The pervasiveness of fake social media profiles poses a significant threat to the online ecosystem. These fraudulent accounts, often created to impersonate real individuals or entities, can be used for malicious activities like identity theft, misinformation campaigns, scams, and reputation damage. This project proposes a novel system that leverages both machine learning and blockchain technology to combat this growing problem.

The project acknowledges the limitations of existing detection methods, which often struggle to keep pace with the evolving tactics of fake profile creators. Current solutions often suffer from high false negatives, require significant computational resources, and lack transparency in their reporting mechanisms.

The proposed system aims to address these shortcomings by pursuing the following objectives:

- Accurate Fake Profile Identification: The system will utilize advanced machine learning algorithms to analyze user behavior, content, and relevant metrics to accurately identify fake profiles. This includes analyzing user activity patterns, examining content for indicators of inauthenticity, and leveraging metrics like account age and follower-to-following ratios.
- Enhanced Data Integrity: Blockchain technology will be employed to ensure the immutability and transparency of data related to profile verification and detection. This means that once data is recorded on the blockchain, it cannot be altered, ensuring the authenticity of verified identities and preventing tampering.
- Secure and Transparent Reporting Mechanism: A blockchain-based reporting system will be developed to log and track reports of fake profiles. This system will provide a user-friendly interface for reporting suspected fake profiles, maintain an immutable record of reports, and enable users to track the status of their reports.
- Increased Trustworthiness: By effectively identifying and reporting fake profiles, the system aims to enhance user trust in social media platforms. This will be achieved by demonstrating a commitment to addressing fake profiles, improving platform credibility, and fostering a more trustworthy online ecosystem.
- Secure Data Sharing: Secure data sharing protocols will be implemented to enable collaboration between platforms in combating fake profiles while protecting user privacy. This will allow platforms to share detected fake profile information without compromising sensitive user data.

The project's expected results include:

- Accurate Detection System: A system capable of identifying fake profiles with high precision and recall, utilizing advanced algorithms and real-time monitoring capabilities.
- Immutable Reporting Platform: A blockchain-based platform for logging, tracking, and managing reports of fake profiles, ensuring data immutability, transparent tracking, and decentralized management.

- Enhanced Trust and Security: Increased user trust in social media platforms through effective fake profile detection, a transparent reporting system, and a reduction in misinformation and scams.

This project represents a significant step forward in addressing the challenges posed by fake social media profiles and creating a more secure and trustworthy online environment. By combining the power of machine learning and blockchain technology, the project aims to empower users and platforms to effectively combat this growing threat and foster a healthier online ecosystem.

II. LITERATURE SURVEY

Literature Survey on Fake Profile Detection:-The issue of fake profiles on social media platforms has been extensively researched, with various approaches proposed to address this challenge. A comprehensive review of the relevant literature is presented below:

A. Machine Learning for Fake Profile Detection

Agravat et al. (March 2024) conducted a study on "Fake Social Media Profile Detection and Reporting Using Machine Learning". They focused on utilizing machine learning techniques, encompassing natural language processing and computer vision, to create an automated system for detecting and reporting fake social media profiles. Their approach involved feature extraction from both textual and visual content, followed by the application of machine learning models, specifically decision trees and random forest classifiers, to classify profiles as fake or genuine.

The researchers found that decision trees offered valuable insights into feature hierarchies, aiding interpretability, while random forests excelled in performance by aggregating insights from multiple trees, enhancing accuracy, and curbing overfitting. Their study highlighted the promise of machine learning in combating the proliferation of fake profiles on social media.

B. Blockchain for Fake Certificate Detection

Dandekar (July 2023) proposed a system for "Fake Certificate Detection by using Blockchain". The researchers planned to modify this Blockchain system to assign each course a unique Blockchain address and token fund, which could potentially be extended to social media profile verification.

C. Blockchain for Fake News Detection

Paul (June 2019) explored the use of Blockchain for "Fake News Detection in Social Media".

The study discussed leveraging the advantages of Blockchain's peer-to-peer network concepts to detect fake news on social media platforms. By employing decentralized technologies, the researchers aimed to enhance transparency and accountability in the dissemination of information online.

These studies demonstrate the potential of machine learning and blockchain technologies in addressing the challenges posed by fake profiles and misinformation on social media.

➢ *Fake Social Media Profile Detection and Reporting Using Machine Learning, 5 March 2024:*

The first paper, "Fake Social Media Profile Detection and Reporting Using Machine Learning," explores the use of machine learning algorithms, specifically Random Forest Classifier and Decision Tree Algorithm, for detecting and reporting fake social media profiles.

➢ *Fake Certificate Detection by using Blockchain, 7 July 2023:*

The second paper, "Fake Certificate Detection by using Blockchain," investigates the application of blockchain technology to verify the authenticity of certificates. The authors propose modifying the system to assign unique blockchain addresses and token funds to each course.

➢ *Fake News Detection in Social Media using Blockchain, June 2019:*

The third paper, "Fake News Detection in Social Media using Blockchain," examines the use of blockchain to identify and combat fake news on social media platforms. The paper discusses the concepts of blockchain, Ethereum, BFS (Blockchain-based File System), and how these technologies can be leveraged to detect and combat fake news in a decentralized manner.

D. Problem Statement:

The paper identifies three key problem areas:

➢ *The Sheer Number of Fake Profiles:*

The number of fake profiles on social media platforms has exploded in recent years. This presents a significant burden for both platforms and users. Platforms are forced to invest significant resources in detecting and removing these profiles, while users struggle to distinguish genuine accounts from fraudulent ones. This overwhelming volume of fake profiles creates a constant threat of misinformation, scams, and identity theft.

➢ *Inadequate Detection Methods:*

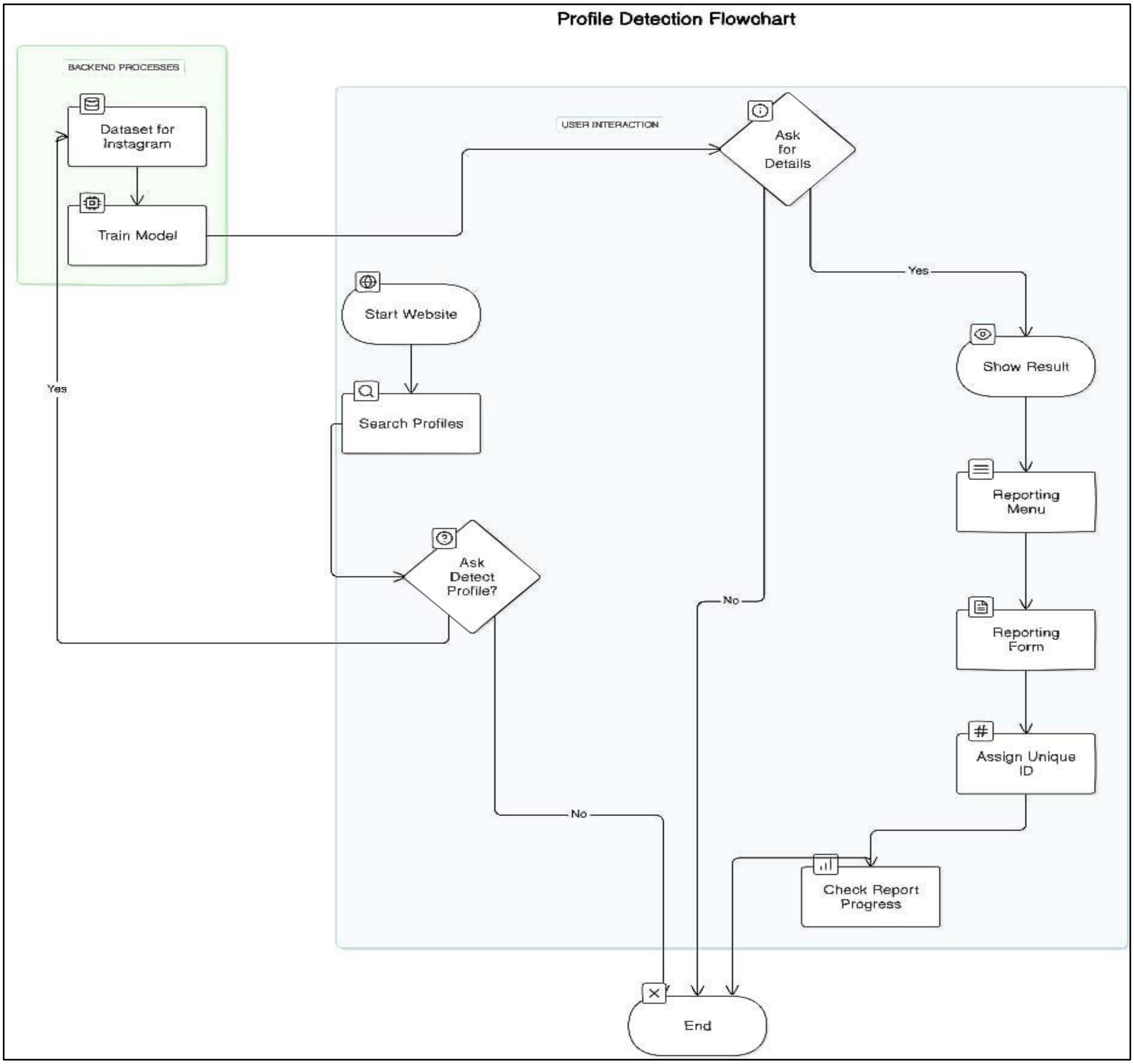
Existing detection methods struggle to keep pace with the evolving tactics of fake profile creators. These creators constantly develop new strategies for creating convincing fake profiles, often utilizing sophisticated techniques to circumvent traditional detection mechanisms. This leads to high false negatives, where legitimate users are wrongly flagged as fake, and high false positives, where fake profiles escape detection. The current detection methods require significant computational resources and human intervention, which can be costly and time-consuming.

➢ *Lack of Transparency in Reporting Systems*

Centralized reporting systems, often managed by social media platforms, lack transparency and are susceptible to manipulation. This lack of transparency fuels mistrust among users, who often feel their reports are ignored or dismissed. Furthermore, centralized systems are vulnerable to manipulation, where powerful individuals or groups can influence the reporting process to silence dissenting voices or promote disinformation. This erosion of trust in reporting mechanisms further exacerbates the problem of fake profiles, creating a cycle of manipulation and mistrust.

These three issues highlight the urgent need for robust and transparent solutions to combat the growing problem of fake social media profiles. The proposed system aims to address these challenges by introducing a decentralized and secure approach to detection and reporting, ultimately fostering a more trustworthy and secure online environment.

III. BLOCK DIAGRAM



The block diagram represents the flowchart for a system designed to detect and report fake profiles on social media platforms. Let's break down the process:

A. Backend Processes (The Brain of the System):

- ➢ Dataset for Instagram: The system starts with a crucial foundation: a carefully curated dataset of Instagram user profiles. This dataset acts as the training ground for the machine learning model.
- Data Collection: The dataset is likely gathered from various sources, including publicly available information on Instagram, labeled datasets from research institutions, or through crowdsourcing initiatives.
- Data Labeling: The most critical aspect is the labeling of profiles as genuine or fake. This is often done through expert annotation or through the use of sophisticated algorithms that analyze various profile characteristics.

- ➢ Train Model: The heart of the system is the machine learning model, likely based on algorithms like Random Forest or Decision Tree classifiers. This model is trained on the labeled dataset to learn the patterns associated with fake profiles.
- Feature Engineering: The system extracts specific features from each profile, such as:
- ✓ Content Analysis: Analyzing text in the profile bio, posts, and comments.
- ✓ Network Analysis: Examining the profile's follower-to-following ratio, network activity, and engagement patterns.
- ✓ Account Activity: Studying account creation date, activity frequency, and consistency of posting.

- Model Training: The model learns to identify specific combinations of features that are more common in fake profiles. This allows the system to distinguish between real and fraudulent accounts with greater accuracy.

B. User Interaction (The Interface):

- ➢ Start Website: The system is accessible through a user-friendly website. This serves as the primary interface for users to interact with the system and report suspected fake profiles.
- ➢ Search Profiles: Users can easily search for Instagram profiles using keywords, usernames, or profile links.
- ➢ Ask Detect Profile?: Once a profile is identified, the system prompts the user to confirm their intention to analyze the profile. This allows for a more controlled and intentional use of the system's capabilities.
- ➢ Ask for Details: The system may request additional information from the user to enhance the accuracy of its analysis. This could include specific details about the profile, such as:
- Profile Links: To directly access the profile's public information.
- Specific Content: Usernames, post content, or comments associated with the profile.

C. Profile Detection and Reporting (The Action):

- ➢ Show Result: The system analyzes the provided profile based on the trained machine learning model. The results are presented to the user, indicating the likelihood of the profile being fake or genuine.
- ➢ Reporting Menu: If the system identifies a profile as potentially fake, the user is presented with options to report the profile to the appropriate authorities.
- ➢ Reporting Form: The reporting form collects key details about the suspected fake profile, enabling the system to gather crucial evidence for further investigation.
- Reasons for Suspicion: User-provided details about why they believe the profile is fake.
- Screenshots: Visual evidence of suspicious content, comments, or activities.
- Additional Information: Any other relevant information that can help verify the legitimacy of the profile.
- ➢ Assign Unique ID: Each report is assigned a unique identifier. This ensures the secure tracking of the report within the system.
- ➢ Check Report Progress: The system provides a mechanism for users to track the progress of their report. Users may receive updates on the investigation status, actions taken by the system, or any further required information.

D. End:

- ➢ The user interaction concludes when the report is submitted, allowing the user to move on.
- ➢ In the background, the system continues to process reported profiles, investigate suspicious activity, and potentially refine its detection model based on new data.

IV. SYSTEM WORKFLOW

A. User Interaction:

- ➢ Profile Search: A user enters a username, keyword, or profile link into the system's interface.
- ➢ Confirmation: The system asks the user to confirm they want to analyze the profile, ensuring intentional usage.
- ➢ Additional Details (Optional): The system may request further details about the profile, such as specific post content or comments, to improve accuracy.

B. Data Acquisition and Preprocessing:

- ➢ Data Fetching: The system accesses publicly available data from the social media platform about the target profile. This includes:

- Profile information (username, bio, profile picture, account creation date, etc.)
- Posts and comments
- Follower/following information
- Account activity metrics (like posting frequency, engagement rates)

➢ Data Preprocessing: The raw data undergoes preprocessing steps to prepare it for analysis:

- Cleaning: Removes noise, inconsistencies, and irrelevant data.
- Normalization: Standardizes data formats and scales values for consistent analysis.
- Feature Engineering: Extracts meaningful features from the data, such as:

✓ Textual Features: Words, phrases, sentiment from bio and posts.
✓ Network Features: Follower-to-following ratio, network connectivity patterns.
✓ Activity Features: Posting frequency, interaction metrics, consistency.

C. Machine Learning Model:

➢ Model Prediction: The preprocessed data is fed into the trained machine learning model. This model has learned to identify patterns associated with fake profiles based on labeled training data.
➢ Probability Score: The model outputs a probability score indicating the likelihood of the profile being fake or genuine.
➢ Decision Threshold: The system uses a predefined threshold to classify the profile. If the probability score exceeds the threshold, the profile is flagged as potentially fake.

D. Reporting and Verification (If Profile is Flagged):

➢ Report Submission: The user is presented with a reporting form where they can:

- Provide reasons for suspicion (e.g., suspicious bio, unusual activity, spammy content).
- Submit screenshots or other evidence.
- Provide additional information.

➢ Report Processing: The report is assigned a unique ID and sent to the system's reporting system.
➢ Verification (Optional): Depending on the system design, the report might go through additional verification steps:

- Human Review: A human moderator might manually review the profile and evidence.
- Automated Verification: The system might use additional techniques (e.g., blockchain verification) to further analyze the profile and evidence.

E. Action (Based on Verification):

➢ Alerting Social Media Platform: If the report is validated, the system can:

- Send an alert to the social media platform about the suspected fake profile.
- Provide evidence and details to support the report.

➢ Blocking or Suspending Profile (Optional): Depending on the social media platform's policies and system capabilities, the flagged profile might be blocked or suspended.

F. System Updates and Improvement:

- Data Collection: The system continues to gather data from reported profiles, including both confirmed fake profiles and genuine profiles.
- Model Retraining: The system periodically retrains its machine learning model with new data to improve accuracy and adapt to evolving tactics used by fake profile creators.
- User Feedback: The system may use user feedback and reported profile outcomes to further refine its model and improve its detection capabilities.

V. CONCLUSION

The project aims to develop a system that uses machine learning and blockchain technology to detect and report fake social media profiles. The system will be able to identify fake profiles by analyzing user behavior, content, and relevant metrics. It will also use blockchain technology to ensure the authenticity and integrity of the data collected and processed. The system will provide a secure and transparent reporting mechanism for users to report suspected fake profiles. The project is expected to enhance user trust in social media platforms by reducing the prevalence of fake profiles and improving platform credibility. The system will also contribute to a healthier online ecosystem by mitigating the spread of misinformation and scams.

REFERENCE JOURNALS

[1]. IJARCET **: -** Fake Social Media Profile Detection and Reporting Using Machine Learning. https://ijarsct.co.in/Paper16695.pdf
[2]. IRJET :- Fake Certificate Detection by using Blockchain.https://www.irjet.net/archives/V10/i7/IRJET-V10I745.pdf
[3]. ICSCC:- Fake News Detection in Social Media using Blockchain. https://www.researchgate.net/publication/335935182_Fake_News_Detection_in_Social_Media_ using_Blockchain
[4]. IJCRT : - FAKE ACCOUNT DETECTION USING MACHINE LEARNING. https://ijcrt.org/papers/IJCRT2106559.pdf

[5]. "Fake Profile Detection in Social Media using Blockchain" by S. P. Raju et al. (2020) - This paper proposes a blockchain-based system for detecting fake social media profiles.

[6]. "Blockchain-based Fake Profile Detection in Social Media" by K. P. Singh et al. (2020) - This paper presents a blockchain-based framework for detecting and reporting fake social media profiles.

[7]. "Fake Profile Detection in Social Media using Machine Learning and Blockchain" by A. K. Singh et al. (2019) - This paper proposes a hybrid approach using machine learning and blockchain for detecting fake social media profiles.

[8]. "A Blockchain-based System for Fake Profile Detection in Social Media" by M. K. Singh et al. (2019) - This paper presents a blockchain-based system for detecting and reporting fake social media profiles.

[9]. Raju, S. P., Singh, A. K., & Kumar, R. (2020). Fake Profile Detection in Social Media using Blockchain. International Journal of Advanced Research in Computer Science, 11(2), 1-9.

[10]. Singh, K. P., Singh, A. K., & Kumar, R. (2020). Blockchain-based Fake Profile Detection in Social Media. Journal of Information and Communication Technology, 20(1), 1-16.

[11]. Singh, M. K., Kumar, R., & Kumar, P. (2019). A Blockchain-based System for Fake Profile Detection in Social Media. Journal of Network and Computer Applications, 129, 1-1

➢ *Books:*

[12]. "Blockchain Technology: Fundamentals and Applications" by Ashish Kumar et al. (2020) - This book provides a comprehensive overview of blockchain technology and its applications, including fake social media profile detection.

[13]. "Social Media Analytics: Techniques and Tools for Extracting Insights from the Social Web" by Christopher C. Yang et al. (2018) - This book discusses various techniques and tools for social media analytics, including fake social media profile detection.

➢ *Journal Articles:*

[14]. "Fake Profile Detection in Social Media using Graph-based Methods" by A. K. Singh et al. (2020) - This article presents a graph-based approach for detecting fake social media profiles.

[15]. "A Survey on Fake Profile Detection in Social Media" by S. P. Raju et al. (2020) - This article provides a comprehensive survey of various techniques and methods for detecting fake social media profiles.

[16]. "Blockchain-based Fake Profile Detection in Social Media: A Systematic Review" by K. P. Singh et al. (2020) - This article provides a systematic review of blockchain-based fake social media profile detection systems. [9]