Fake Instagram Profile Identification and Classification using Machine Learning

Aniruddha Lalge, Apurv Badave,
Nikhil Elajale, Pankaj Godara
Department of Computer Engineering,
Smt. Kashibai Navale College of Engineering
Pune- 41, India

Prof. Priyanka Kinage
Priyanka.kinage_skncoe@sinhgad.edu
Department of Computer Engineering
Smt. Kashibai Navale College of Engineering
Pune- 41, India

***Abstract*— Social media platforms have become integral to modern communication, enabling users to connect, share, and engage in various activities. However, the rise of fake profiles on platforms like, Instagram possess significant challenges related to user privacy, security, and trust. This work presents a novel approach to identify and classify fake Instagram profiles using machine learning techniques. The findings of this research contribute to the ongoing efforts to combat the proliferation of fake profiles on Instagram and other social media platforms. By leveraging machine learning techniques and a comprehensive feature set, the proposed model demonstrates promising results in identifying and classifying fake profiles, thereby promoting a safer and more trustworthy online environment. This research opens avenues for further exploration, including the integration of real-time data streams and the adaptation of the model to other social media platforms.**

Keywords: Profile identification, User authentication, Data preprocessing, Model training, Online security, Machine learning**.**

I. INTRODUCTION

Fake Instagram profiles can range from automated bots posting spam to sophisticated imposters attempting to deceive genuine users for financial gain, social manipulation, or other illicit activities. Traditional methods of manual inspection and reporting are insufficient to handle the sheer volume of profiles and interactions, necessitating the use of advanced technological solutions. Machine learning has emerged as a powerful tool in addressing the issue of fake profiles on social media platforms. By harnessing the computational power of machine learning algorithms, it is possible to automatically identify and classify fake profiles based on distinctive patterns and characteristics. The combination of the growing influence of social media, the challenges posed by fake profiles, and the advancements in machine learning techniques has led to the development of solutions aimed at identifying and classifying these profiles. This research addresses the need for a safer and more trustworthy online environment by proposing a comprehensive approach to tackle the issue of fake Instagram profiles using machine learning.

II. PROPOSED SYSTEM

Creating a robust algorithm for detecting fake profiles on Instagram is a challenging but important task. Instagram, like many other social media platforms, is plagued by fake profiles that engage in spam, fraud, or other malicious activities. Here's a proposed system for an Instagram fake profile detection algorithm:

1. Data Collection:
Gather a large dataset of Instagram profiles, including both genuine and fake profiles, with a variety of characteristics.
2. Feature Extraction:
Extract relevant features from user profiles. These features can include:

- Profile picture analysis: Check for low-quality images, inconsistencies, or reused images.
- Activity patterns: Analyze the frequency of posts, likes, comments, and followers.
- Bio information: Look for inconsistencies, unusual characters, or common patterns used by spammers.
- Follower-to-following ratio: Check for extreme imbalances.
- Post content analysis: Analyze the content of posts for spammy keywords, URLs, or trends.

3. Machine Learning Model:
Develop a machine learning model to classify profiles as genuine or fake. Consider using techniques like:
Supervised learning with labeled data.
Deep learning models, such as neural networks, to capture complex patterns.
Ensemble techniques like Random Forest or Gradient Boosting for better accuracy.
4. Training and Validation:
Split the dataset into training, validation, and test sets.Train the model on the training data, fine-tuning the hyper parameters.
Validate the model's performance on the validation set and iterate if necessary.
5. Real-time Monitoring:
Implement the algorithm to work in real-time on Instagram profiles. Continuously monitor user activity, profiles, and interactions for potential fake behavior.

A. Support Vector Machine :

Support Vector Machines (SVMs) are a popular type of supervised machine learning algorithm used for classification and regression tasks. SVMs can perform both linear and non-linear classification.

Support Vector Machine Working Steps:

Step 1: Load the important libraries.
Step 2: Import dataset and extract the X variables and Y separately.
Step 3: Divide the dataset into train and test.
Step 4: Initializing the SVM classifier model.
Step 5: Fitting the SVM classifier model.
Step 6: Coming up with predictions.
Step 7: Evaluating model's performance

B. Random Forest :

Random Forest is an ensemble learning method that is widely used for both classification and regression tasks in machine learning. Random Forest is a learning technique that combines the predictions of multiple decision trees to improve the overall predictive accuracy and robustness. Random Forest Working Steps:

Step 1: Importing and processing the data.
Step 2: Training the random forest classifier.
Step 3: Testing the prediction accuracy.
Step 4: Visualizing the results of the classifier.

III. SYSTEM DESIGN

Designing a system architecture for an Instagram fake profile detection algorithm involves several components, each with specific functions. Here are the components of the architecture:

1. Data Collection Layer
2. Preprocessing Layer
3. Feature Extraction
4. Classification Algorithms (Machine Learning Algorithms)
5. Implementation

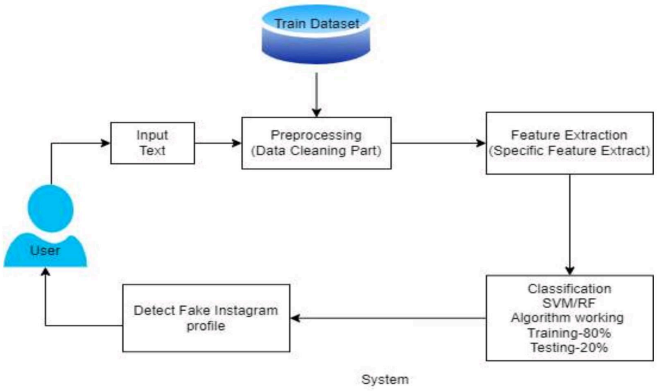


Fig. I System Architecture

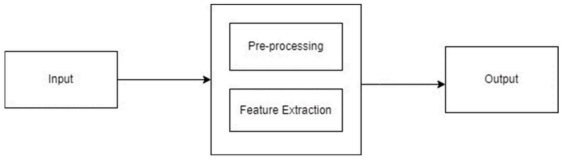
This system architecture provides a comprehensive view of how different components work together to detect fake profiles on Instagram. The key to success is the continuous refinement of the machine learning model, real-time monitoring, and the ability to adapt to evolving strategies used by malicious actors.

A. Data Flow Diagram :

In Data Flow Diagram, we show that flow of data in our system. In DFD0 we show that base DFD in which rectangle represent input as well as output and circle shows our system, In DFD1 we show actual input and actual output of system input of our system is text or image and output is rumor detected likewise in DFD 2 we present operation of user as well as admin.

DFD0 :

DFD1 :



DFD2:

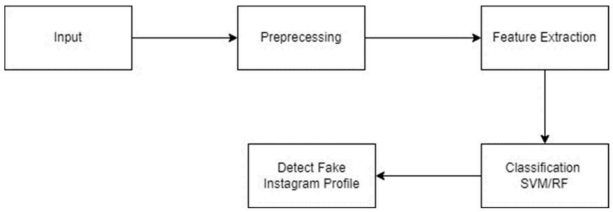


Fig. II Data Flow Diagram

B. Class Diagram:

A Class diagram is a type of UML (Unified Modeling Language) diagram used to illustrate the structure of a system by showing the classes of the system, their attributes, methods, and the relationships among these classes.

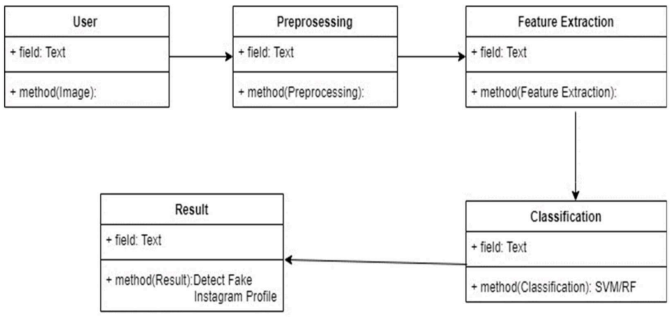


Fig. III Class Diagram

Class represents a blueprint for objects, including its name and structure. Class diagrams provide a visual overview of a system's structure, aiding in design, understanding relationships between classes, and communicating the system's architecture among stakeholders, including developers, designers, and stakeholders. They serve as a foundation for implementing the software system in an object-oriented approach.

IV. ANALYSIS MODEL: SDLC MODEL

In traditional software development, "analysis models" typically refer to the models created during the requirements analysis phase of the software development life cycle (SDLC). These models include use case diagrams, data flow diagrams, entity-relationship diagrams, and other visual representations used to capture and understand the software requirements. The software development cycle is a combination of different phases such as designing, implementing and deploying the project. The SDLC model for the project development can be understood using the following steps. The chosen SDLC model is the waterfall model which is easy to follow and fits bests for the implementation of this project.

- Requirements Analysis: At this stage, the business requirements, definitions of use cases are studied and respective documentations are generated.
- Design: In this stage, the designs of the data models will be defined and different data preparation and analysis will be carried out.
- Implementation: The actual development of the model will be carried out in this stage. Based on the data model designs and requirements from previous stages, appropriate algorithms, mathematical models and design patterns will be used to develop the agent's back-end and front-end components.
- Testing: The developed model based on the previous stages will be tested in this stage. Various validation tests will be carried out over the trained model. Deployment: After the model is validated for its accuracy scores its ready to be deployed or used in simulated scenarios.
- Maintenance: During the use of the developed solution various inputs/scenarios will been countered by the model which might affect the models overall accuracy Or with passing time the model might not fit the new business requirements. Thus, the model must be maintained often to keep its desired state of operation.

V. FEASIBILITY AND SCOPE

Fake profiles often exhibit certain characteristics that can be used for identification. These characteristics include limited or generic profile pictures, a small number of followers, a high number of followings, erratic posting patterns, and low engagement on posts. Analyzing the content posted by a user can also provide insights. Fake profiles may frequently repost or share content from other accounts, use generic captions, or engage in behaviors like excessive use of hashtags. Examining the user's behavior on the platform can be informative. Fake profiles may engage in activities such as mass following and unfollowing, liking and commenting on numerous posts in a short time, or sending unsolicited messages. Assess the financial resources required for data collection, machine learning model development, real-time monitoring infrastructure, and system maintenance. Evaluate the potential costs associated with the development and maintenance of the system against the benefits of reducing the impact of fake profiles on Instagram.

VI. BACK UP PLAN

If collecting a sufficient amount of labeled data is challenging, consider data augmentation techniques to create synthetic data that can help improve model performance. Implement a mechanism for users to report suspicious profiles. Combine user reports with automated detection for more accurate results. Create ensemble models by combining multiple machine learning algorithms to increase accuracy and robustness. Work closely with legal teams and consider implementing stricter policies for account creation and verification. Ensure that your model respects user privacy and ethical guidelines, especially when handling user data and profile information.

VII. ADVANTAGES

- Users can have a more positive and trustworthy experience on Instagram when fake profiles are identified and removed.
- Fake profiles can be used for various malicious purposes, including harassment, spamming, and spreading misinformation.
- Identifying and removing such profiles, Instagram can contribute to reducing the spread of misinformation on the platform.
- Fake profiles can be used to impersonate legitimate users or organizations for phishing scams.
- Social media algorithms often prioritize content and profiles based on user engagement.
- Many jurisdictions have regulations that require social media platforms to take action against fake profiles and to protect user data. Implementing a fake profile detection system helps ensure legal compliance.
- Fake profiles can be used to scrape and misuse user data. Detecting and eliminating these profiles helps protect user data from unauthorized access.
- Platforms with fewer fake profiles and a better reputation can attract more legitimate advertisers and content creators, leading to increased revenue opportunities.

VIII. LIMITATIONS

- Malicious actors are continually evolving their tactics to create more sophisticated fake profiles that can evade detection algorithms.
- Analyzing user profiles and behavior for the purpose of fake profile detection raises concerns about user privacy.
- Real profiles significantly outnumber fake profiles on Instagram.
- New types of fake profiles and malicious behavior emerge regularly, making it difficult for detection algorithms to keep up with evolving threats.

- Some legitimate users may employ automation tools for benign purposes, such as social media management.
- Smaller social media platforms or startups may lack the resources to develop and maintain robust fake profile detection systems.

IX. CONCLUSION

The research on ”Fake Instagram Profile Identification and Classification using Machine Learning” presents a comprehensive approach to tackle the persistent issue of fake profiles on social media platforms, with a specific focus on Instagram. By leveraging the power of machine learning techniques, this research contributes to creating a safer and more trustworthy online environment for users, bolstering user confidence, and upholding the integrity of social media community. The research’s outcomes extend beyond the realm of academia, impacting the lives of individuals, businesses, and society as a whole. As social media continues to shape the digital landscape, the work presented here contributes to building a foundation of trust and authenticity, reinforcing the positive potential of online interactions and collaborations.

X. REFERENCES

[1] Zainab Agha, Neeraj Chatlani, Afsaneh Razi, and Pamela Wisniewski. 2020. Towards conducting responsible research with teens and parents regarding online risks. In Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems. 1–8.

[2] Shiza Ali, Afsaneh Razi, Seunghyun Kim, Ashwaq Alsoubai, Joshua Gracie, Munmun De Choudhury, Pamela J Wisniewski, and Gianluca Stringhini. 2022. Understanding the Digital Lives of Youth: Analyzing Media Shared within Safe Versus Unsafe Private Conversations on Instagram. (2022), 1–14.

[3] Detect fake profiles on social media network https://telanganatoday.com/detect-fake-profiles-on-social-medianetworks

[4] S. M. Din, R. Ramli and A. A. Bakar, "A Review on Trust Factors Affecting purchase Intention on Instagram", 2018 IEEE Conference on Application Information and Network Security (AINS), 2018

[5]. S.C. Boerman, "The effects of the standardized Instagram disclosure for micro- and meso-influencers", Computers in Human Behavior, vol. 103, pp. 199-207, 2020.

[6]. S. Sheikhi, An Efficient Method for Detection of Fake Accounts on the Instagram Platform, 2020.

[7]. J. Kang and L. Wei, "Let me be at my funniest: Instagram users' motivations for using Finsta (a.k.a. fake Instagram)", The Social Science Journal, 2019.

[8]. M. Mondal, L. A. Silva and F. Benevenuto, "A Measurement Study of Hate Speech in Social Media", Proceedings of the 28th ACM Conference on Hypertext and Social Media - HT '17, 2017.

[9]. B. Mathew, R. Dutt, P. Goyal and A. Mukherjee, "Spread of Hate Speech in Online Social Media", Proceedings of the 10th ACM Conference on Web Science.

[10]. H. Hilal Bashir and S. A. Bhat, "Effects of Social Media on Mental Health: A Review", The International Journal of Indian Psychology, vol. 4, no. 3, 2017

[11]. Aleksei Romanov, Alexander Semenov, Oleksiy Mazhelis and Jari Veijalainen.2017.”Detection of Fake Profiles in Social Media”. In 13th International Conference on Web Information Systems and Technologies.

[12]. Indira Sen,Anupama Aggarwal,Shiven Mian.2018.”Worth its Weight in Likes: Towards Detecting Fake Likes on Instagram”. In ACM International Conference on Information and Knowledge Management.Nazir, Atif, Saqib Raza, Chen-Nee Chuah, Burkhard Schipper, and C. A. Davis. “Ghostbusting Facebook: Detecting and Characterizing Phantom Profiles in Online Social Gaming Applications.” In WOSN. 2010.

[13]. Nambouri Sravya, Chavana Sai praneetha, S. Saraswathi,” Identify the Human or Bots Twitter Data using Machine Learning Algorithms”, International Research Journal of Engineering and Technology (IRJET), Volume: 06 Issue: 03 — Mar 2019 www.irjet.net, e-ISSN: 2395-0056, p- ISSN: 2395-0072.

[14]. M. Smruthi, N. Harini,” A Hybrid Scheme for Detecting Fake Accounts in Facebook”, International Journal of Recent Technology and Engineering (IJRTE) ISSN: 2277-3878, Volume-7, Issue-5S3, February 2019.

[15]. Nazir, Atif, Saqib Raza, Chen-Nee Chuah, Burkhard Schipper, and C. A. Davis. “Ghostbusting Facebook: Detecting and Characterizing Phantom Profiles in Online Social Gaming Applications.” In WOSN. 2010.

[16]. Rao, P. S., J. Gyani, and G. Narsimha. ”Fake profiles identification in online social networks using machine learning and NLP.” Int. J. Appl. Eng. Res 13.6 (2018): 973-4562.

[17]. Raturi, Rohit. ”Machine learning implementation for identifying fake accounts in social network.” International Journal of Pure and Applied Mathematics 118.20 (2018): 4785-4797. J. Wang, “Fundamentals of erbium-doped fibre amplifiers arrays (Periodical style—Submitted for publication),” IEEE J. Quantum Electron., submitted for publication.

[18]. M. Mohammadrezaei, M. E. Shiri, and A. M. Rahmani, “Identifying fake accounts on social networks based on graph analysis and classification algorithms,” Security and Communication Networks, vol. 2018, 2018.

[19]. Ala’M, Al-Zoubi, Ja’far Alqatawna, and Hossam Faris. ”Spam profile detection in social networks based on public features.” 2017 8th International Conference on information and Communication Systems (ICICS). IEEE, 2017.