IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.53
Volume 4, Issue 5, March 2024

Fake Social Media Profile Detection and Reporting Using Machine Learning

Aniket Agravat, Umang Makwana, Sahil Mehta, Devashish Mondal, Sushant Gawade
Department of Artificial Intelligence and Machine Learning
Universal College of Engineering, Mumbai, Maharashtra, India

Abstract: *Our research focuses on utilizing machine learning techniques, encompassing natural language processing and computer vision, to create an automated system for the detection and reporting of fake social media profiles across various platforms. Our approach involves feature extraction from both textual and visual content, followed by the application of machine learning models to classify profiles as fake or genuine. This system operates in real-time, monitoring user activity and promptly flagging suspicious profiles for user- initiated reporting. By combining the power of machine learning with cross-platform compatibility and user feedback, our solution aims to enhance online safety by swiftly identifying and addressing fraudulent social media profiles, thus fostering more secure and trustworthy online communities.*

Keywords: Fake profiles, Machine learning, Natural Language Processing, Fraudulent Social media accounts

I. INTRODUCTION

A website known as a "social networking site" is one where users may connect with friends, make updates, and find new people who have similar interests. Each user has a profile on the website. Users can communicate with one another using Web 2.0 technologies in these online social networks [1]. The utilization of social networking sites is expanding quickly and affecting how individuals interact with one another. Online communities bring together people with like interests and make it easy for users to find new friends. The main benefit of Internet social networking is that it allows users to easily connect with people and communicate better. This has provided new avenues for potential attacks such as fake identities, disinformation, and more [3]. Researchers are working to determine the impact these online social networks have on people. There is much more to media than just how many people use it. This suggests that the number of fake accounts has grown throughout the past years [4]. A website known as a "social networking site" is one where users may connect with friends, make updates, and find new people who have similar interests. Each user has a profile on the website. Users can communicate with one another using Web 2.0 technologies in these online social networks [1]. The utilization of social networking sites is expanding quickly and affecting how individuals interact with one another. Online communities bring together people with like interests and make it easy for users to find new friends. The main benefit of Internet social networking is that it allows users to easily connect with people and communicate better. This has provided new avenues for potential attacks such as fake identities, disinformation, and more [3]. Researchers are working to determine the impact these online social networks have on people. There is much more to media than just how many people use it. This suggests that the number of fake accounts has grown throughout the past years [4].

Detecting fake profiles on social media is vital for maintaining trust, protecting users from harm, preserving, authenticity, and ensuring effective advertising. Fake profiles can spread misinformation, scam users, and distort user data, impacting the credibility and security of the platform. By identifying and removing fake profiles, social media platforms can uphold user trust, safeguard privacy, and maintain the integrity of interactions and advertising efforts.

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.53
Volume 4, Issue 5, March 2024

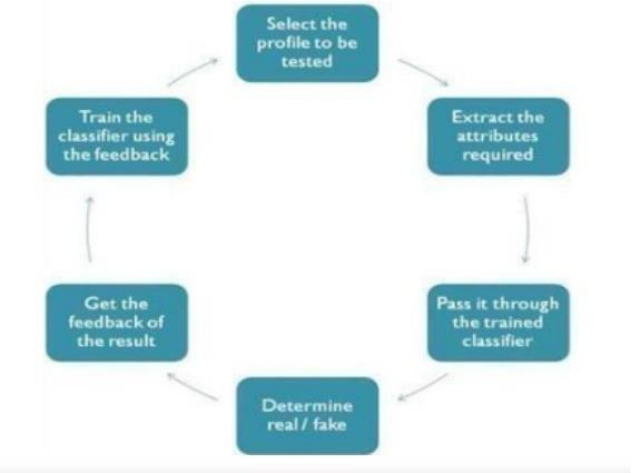

Fig. 1: Detection Process

Fig. 2: Flowchart

II. LITERATURE REVIEW

In this paper, the survey provides us with a comprehensive review of important techniques for fake profile detection in OSNs. This paper focuses on the state of the art for detecting Sybil or fake accounts. This paper uses Machine Learning and Graph Base Classification. The open issue in the domain of fake profile detection is stated. Hadoop and Spark will definitely be the part of solution for large amounts data. This paper reports on a study that is focused on detecting fake accounts created by humans, as opposed to those created by bots. The machine learning models were trained to use engineered features without relying on behavioral data. This made it possible for these machine learning models to be trained on very little data, compared to when behavioral data is included. Future work will investigate the enrichment of the feature. set used in the research for this paper by engineering features from the social sciences knowledge domain.

The administrator controls the entire application, admin should train the model and make it ready to process. If model needs to get further trained, that process would be handled by the admin. This software has been computed successfully. A detection method has been proposed which can find both fake and clone Twitter profiles. The problem can be further extended to Natural Language Processing Techniques. In this paper our proposed solution for this problem is to train 70% of our profiles data on machine learning algorithms using Spark ML lib and then we will test remaining 30% data to find accuracy and predictions. The prediction model will be depending on steps such as reading data set from CSV feature engineering, training data using Random Forest, plotting learning curve, plotting confusion matrix, plotting ROC curve. A new classification algorithm was proposed to improve detecting fake accounts on social networks, where the SVM trained model decision values were used to train a NN model, and SVM testing decision values were used to test the NN model.

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.53
Volume 4, Issue 5, March 2024

III. METHODOLOGY

3.1 System Block Diagram

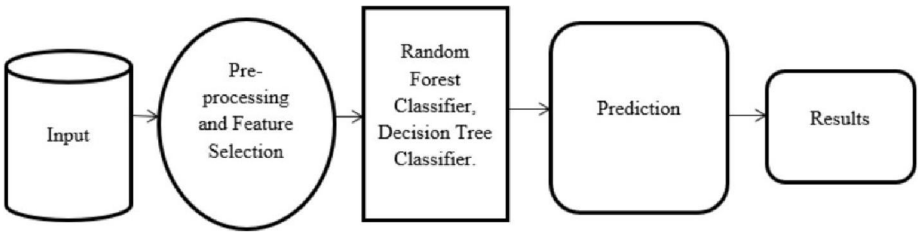

Fig. 3: Block diagram

The system architecture for the college project, "Fake Social Media Profile Detection and Reporting System," is designed to detect and report fake social media profiles effectively. At its core, the architecture consists of a User Interface (UI) through which users can interact with the system. The Frontend and Backend components handle the user interface and user requests, respectively, facilitating the reporting process. User reports, including evidence and profile details, are managed by the Reporting System, which also triggers real-time alerts when a profile is marked as suspicious. The Machine Learning Models are responsible for analyzing profiles, classifying them as genuine or fake based on extracted features and user-supplied evidence. The Database securely stores detected fake profiles and user reports, ensuring data privacy and regulatory compliance. Additionally, an Administrator Panel offers administrators tools for profile management. The system includes a Performance Evaluation component to assess its accuracy and efficiency, informing future improvements continuously. This architecture provides a structured framework for the college project, offering a clear overview of the system's components and their interactions.

3.2 Hardware & Software Details

Table 1: Hardware Requirements for Development of Project:

Sr. No.	Hardware	Specification
1.	Computer/Laptop	Installed windows
2.	Internet connection	2 mbps
3.	RAM	4 GB

Table 2: Software Requirements for the Development of the Project:

Sr. No.	Software	Specification
1.	Python	3.12.2
2.	VS Code	Version 1.87
3.	Python Libraries	Pandas, NumPy, Matplotlib

3.3 Algorithms

A. Random Forest Classifier

A random forest is a powerful ensemble learning technique used for classification and regression tasks. It operates by constructing multiple decision trees based on random subsets of the training data and then combining their predictions to improve accuracy and reduce overfitting.

Here's how a random forest typically works:

1. Randomly select a subset of data points (with replacement) from the training set.
2. Build a decision tree using the selected subset.
3. Repeat steps 1 and 2 multiple times to create a forest of decision trees.
4. When making predictions for new data points, each decision tree in the forest provides a prediction.
5. The final prediction is determined by combining the individual predictions through a majority voting process for classification tasks or averaging for regression tasks.

This approach helps random forests to generalize well to new data and improve predictive performance compared to individual decision trees.

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.53
Volume 4, Issue 5, March 2024

B. Decision Tree Algorithm

A decision tree is a versatile supervised learning algorithm used for classification and regression tasks. It adopts a hierarchical structure comprising a root node, branches, internal nodes, and leaf nodes, offering intuitive models for analysis.

Here's a breakdown of how the decision tree algorithm functions:

1. Initialization: The algorithm commences at the root node, which encompasses the entire dataset.
2. Feature Selection: It identifies the most crucial feature or question that efficiently divides the data into distinct groups, akin to choosing a path at a branching point.
3. Branching: Based on the chosen feature, the algorithm creates new branches by partitioning the data into smaller subsets, representing alternative paths through the tree.
4. Iterative Process: This process of asking questions and splitting data persists at each node until reaching the terminal leaf nodes, which yield the final predicted outcomes or classifications.

Assumptions and Considerations:

Decision trees rely on several assumptions and considerations to construct effective models, influencing their performance. Here's a rundown of these key assumptions:

1. Binary Splits: Decision trees typically execute binary splits, dividing data into two subsets based on a single feature or condition, assuming each decision is binary.
2. Homogeneity: The objective is to establish homogeneous subgroups within each node, ensuring that samples within a node are similar concerning the target variable to delineate clear decision boundaries.
3. Top-Down Greedy Approach: Decision trees employ a top-down, greedy approach, selecting splits to maximize information gain or minimize impurity at each node, potentially leading to locally optimal rather than globally optimal trees.
4. Handling Features: They accommodate both categorical and numerical features, necessitating different splitting strategies for each type.
5. Overfitting: Decision trees are susceptible to overfitting, capturing noise in data, mitigated through techniques such as pruning and setting appropriate stopping criteria.
6. Impurity Measures: Evaluation of splits relies on impurity measures like Gini impurity or entropy to gauge how effectively splits segregate classes, impacting tree construction.
7. Handling Missing Values: Assumption is made that there are no missing values in the dataset or that missing values have been adequately managed through imputation or similar methods.
8. Equal Feature Importance: Unless feature scaling or weighting is applied, decision trees may assume equal importance for all features.
9. Outliers Sensitivity: They are sensitive to outliers, necessitating preprocessing or robust methods to address extreme values effectively.
10. Sample Size Sensitivity: Balancing sample size and tree depth is crucial to avoid overfitting in small datasets or overly complex trees in large datasets.

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.53
Volume 4, Issue 5, March 2024

IV. RESULTS

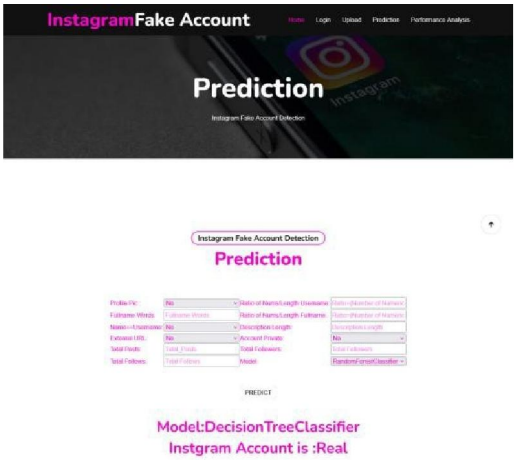

Fig. 4.1: OUTPUT for Real ID

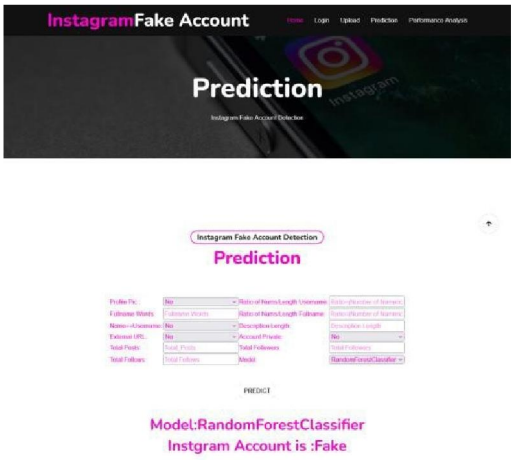

Fig. 4.2: OUTPUT for Fake ID

V. CONCLUSION

In our project, we delved into detecting and reporting fake social media profiles using machine learning, with a focus on decision tree and random forest classifiers. Through feature engineering and model training, we developed robust classifiers adept at spotting patterns indicative of fraudulent behavior. 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. Our study highlights the promise of machine learning in combating the proliferation of fake profiles on social media. By leveraging these algorithms, users and platform administrators gain tools for early detection and prompt reporting of suspicious accounts, crucial for upholding the integrity of online communities. Future research avenues include exploring advanced ensemble methods, deep learning architectures, and integrating contextual data to bolster detection capabilities. Moreover, there's a need for scalable, automated systems for real-time monitoring and response to evolving tactics of malicious actors.Overall, our work contributes to the discourse on employing machine learning to counter online deception, emphasizing the significance of interdisciplinary collaboration among data scientists, social scientists, and platform stakeholders in tackling this pervasive issue.

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.53
Volume 4, Issue 5, March 2024

REFERENCES

[1] E. Karunakar, V. D. R. Pavani, T. N. I. Priya, M. V. Sri, and K. Tiruvalluru, "Ensemble fake profile detection using machine learning (ML)," J. Inf. Comput. Sci., vol. 10, pp. 1071–1077, 2020.
[2] P. Wanda and H. J. Jie, "Deep profile: utilising dynamic search to identify phoney profiles in online social networks CNN" J. Inf. Secur. Appl., vol. 52, pp. 1–13, 2020.
[3] P. K. Roy, J. P. Singh, and S. Banerjee, "Deep learning to filter SMS spam," Future Gener. Comput. Syst., vol. 102, pp. 524–533, 2020.
[4] R. Kaur, S. Singh, and H. Kumar, "A modern overview of several countermeasures for the rise of spam and compromised accounts in online social networks," J. Netw. Comput. Appl., vol. 112, pp. 53– 88, 2018.
[5] G. Suarez-Tangil, M. Edwards, C. Peersman, G. Stringhini, A. Rashid, and M. Whitty, "Automatically dismantling online dating fraud," IEEE Trans. Inf. Forensics Secur., vol. 15, pp. 1128–1137, 2020.
[6] K. Thomas, C. Grier, D. Song, and V. Paxson, ''Suspended accounts in retrospect: An analysis of Twitter spam,in Proc. ACM SIGCOMM Conf. internet Meas. Conf., 2011, pp. 243–258.
[7] Saeed Abu-Nimeh, T. M. Chen, and O. Alzubi, "Malicious and Spam Posts in Online Social Networks," Computer, vol.44, no.9, IEEE2011, pp.23–28.
[8] B. Viswanath et al., ''Towards detecting anomalous user behavior in online social networks,'' in Proc. Usenix Secur., vol. 14. 2014, pp. 223–238.
[9] R. Kaur and S. Singh, "A survey of data mining and social network analysis based anomaly detection techniques," Egyptian informatics journal, vol. 17, no. 2, pp. 199–216, 2016.
[10] S.-T. Sun, Y. Boshmaf, K. Hawkey, and K. Beznosov, "A billion keys, but few locks: the crisis of web single sign-on," in Proceedings of the 2010 New Security Paradigms Workshop. ACM, 2010, pp. 61–7