International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 08 Issue: 11 | Nov - 2024 SJIF Rating: 8.448 ISSN: 2582-3930

# SOCIAL MEDIA FAKE ACCOUNT IDENTIFICATION USING ML

Shraddhesh Dalvi, Shubham Gholap, Piyush Jadhav, Prof. Pallavi Baviskar

[1,2,3,4]*Undergraduate Students Computer Engineering Department, Sandip Institute of Engineering and Management, Nashik, India*
[5]*Professor, Computer Engineering Department, Sandip Institute of Engineering and Management, Nashik, India*

---

**Abstract -** The proliferation of social media platforms has led to an increase in the creation of fake accounts. These accounts are used for various malicious activities, such as spreading false information, phishing, and identity theft. As a result, there is a growing need for effective methods to identify and eliminate fake accounts. This paper proposes a machine learning-based approach for social media fake account identification. This paper proposes a machine learning-based approach to identify fake accounts on social media platforms. Our method leverages a combination of feature extraction techniques and supervised learning algorithms to classify accounts as genuine or fake. We collect a large dataset of labeled social media accounts, including both genuine and fake profiles, and extract features from various sources such as profile information, network behavior, and content analysis. We experiment with multiple machine learning models, including Support vector machines(SVM),K-Nearest Neighbors Algorithm(KNN),Random forest, Logistic Regression & Artificial Neural Network(ANN) to evaluate their performance in identifying fake accounts. Our proposed method has significant implications for social media platform operators, policymakers, and researchers seeking to combat fake accounts and maintain online trust. The approach can be integrated with existing social media moderation tools to enhance the accuracy and efficiency of fake account identification.

***Key Words*:** Support vector machines(SVM),K-Nearest Neighbors Algorithm(KNN),Random forest, Logistic Regression & Artificial Neural Network(ANN),Python**.**

## 1.INTRODUCTION

A project on Social Media Fake Account Identification can leverage machine learning algorithms such as Support Vector Machines (SVM), K-Nearest Neighbors (KNN), Random Forest, Logistic Regression, and Artificial Neural Networks (ANN) to accurately detect fake accounts across social platforms. The approach involves extracting profile features (such as account age, friend connections, post frequency, and engagement patterns) and behavioral metrics to create a dataset representing both real and fake accounts. By training each of these algorithms on this dataset, the system can classify accounts based on patterns typical of genuine vs. fake users. The project can explore the comparative performance of each algorithm, aiming to optimize detection accuracy and speed, and potentially combine multiple models in an ensemble method for robust classification. This model could serve as a backend system for social media companies to proactively identify and flag fake accounts, enhancing user experience and platform security.

The motivation for using machine learning (ML) algorithms, including Support Vector Machines (SVM), K-Nearest Neighbors (KNN), Random Forest, Logistic Regression, and Artificial Neural Networks (ANN), in social media fake account detection lies in the ris- ing misuse of social platforms by malicious entities. Fake accounts are commonly used to spread misinformation, influence public opinion, perpetrate fraud, and invade user privacy. Traditional detection methods often fall short due to the adaptability of fake accounts, which continually evolve to avoid detection. ML algorithms provide a dynamic, scalable solution, enabling platforms to respond to these sophisticated tactics. Each algo- rithm offers unique strengths that enhance detection accuracy. SVM excels in managing high-dimensional data, which is essential for analyzing the multitude of features that characterize fake accounts. Its strong decision boundaries make it effective at distinguishing genuine accounts from fake ones. KNN, with its simplicity, works well when analyzing labeled data by identifying similarities between new accounts and previously flagged ones, revealing anomalies based on proximity. Random Forest, an ensemble method, combines multiple decision trees, providing robustness against overfitting while handling diverse features like profile metadata and user interaction patterns. Logistic Regression, known for interpretability, allows platforms to analyze the probability of an account being fake based on individual features, offering insights into key indicators of fake profiles. ANN, capable of capturing non-linear data relationships, adapts to evolving patterns of fake ac- counts, making it powerful in identifying sophisticated fraudulent behavior. By combining these algorithms, social media platforms can create a comprehensive detection system that adapts to the

evolving landscape of fake accounts. This multi-faceted approach bolsters platform security, enhances user trust, and ensures a safer online environment for authentic users. In the face of increasingly sophisticated fake accounts, ML-based methods are essential for maintaining social media integrity and combating misuse.

## 2. PROBLEM STATEMENT

The rise of fake accounts on social media platforms poses significant challenges to user privacy, platform integrity, and public trust. These accounts are often used for malicious purposes such as spreading misinformation, phishing, and manipulating public opinion. Traditional methods of detecting fake accounts are insufficient due to the large scale and complexity of social media data. The problem, therefore, is to develop an effective and scalable solution for identifying fake social media accounts using machine learning (ML) algorithms. This study aims to apply and compare the performance of five widely-used ML algorithms—Support Vector Machines (SVM), K-Nearest Neighbors (KNN), Random Forest, Logistic Regression, and Artificial Neural Networks (ANN)—to determine the most accurate and efficient approach for detecting fake accounts based on user behavior, profile features, and activity patterns.

## 3. OBJECTIVE

The objective of this research is to develop a machine learning-based approach for identifying fake social media accounts. With the increasing prevalence of fake accounts on social media platforms, there is a growing need for effective methods to distinguish between real and fake profiles.

Our proposed method involves a multi-step process that combines data collection, feature engineering, model selection and training, and model deployment and evaluation. By leveraging the power of machine learning algorithms, we aim to accurately distinguish between real and fake social media accounts and mitigate the negative impacts of fake accounts on social media platforms.

The ultimate goal is to provide a reliable and scalable solution for social media com- panies to combat the issue of fake accounts and improve the overall user experience.

We will focus on three popular social media platforms, namely Facebook, Twit- ter, and Instagram, to collect data and train our models. However, the proposed methodology can be applied to other social media platforms as well.

We will collect a large dataset of social media profiles, both real and fake, from these platforms. The dataset will include user behavior, network structure, content anal- ysis, and account metadata. We will also consider factors such as user engagement, account age, and activity patterns to differentiate between real and fake accounts.

We will extract various features from the preprocessed data using techniques such as text analysis, graph theory, and machine learning algorithms. These features will include user behavior patterns, network structure metrics, content analysis features, and account metadata features.

We will experiment with different machine learning algorithms such as logistic regression, support vector machines (SVM), random forests, and neural networks to find the best-performing model for our task. We will also consider ensemble methods such as stacking and boosting to improve the performance of our models.

We will evaluate the performance of our models using various metrics such as accuracy, precision, recall, and F1 score. We will also consider factors such as computational efficiency and scalability when selecting our final model.

## 4. LITERATURE REVIEW

In their Paper by Aksar, Firdaus, and Pasha (2023), titled ”Virtual vs. Real Self: Gendered Presentation and Everyday Performance of Virtual Selfhood A Case Study of Pakistan,” the authors examine how gender influences online self-presentation, highlighting the distinction between virtual and real identities, particularly in the context of Pakistani society. The study offers valuable insights into how women in a gendered cultural environment curate their virtual selves to navigate societal expectations while exploring 4 Rohini Ashok Gamane and Prof. V D dabhade personal freedom online. Its merit lies in shedding light on the challenges and opportunities women face in digital spaces, contributing to the understanding of gendered online behavior in underrepresented regions. However, a key limitation is its focus on a specific cultural context, which may limit the generalizability of the findings to other societies. while the study effectively captures women’s experiences, it under develops the analysis of male virtual selfhood, leaving a gap in understanding the broader spectrum of gendered identity performance online.[1]

In their Paper, ”Machine Learning-Based Social Media Bot Detection: A Comprehensive Literature Review,” M. Aljabri, R. Zagrouba, A. Shaahid, F. Alnasser, A.

Saleh, and D. M. Alomari (2023) provide a thorough examination of various machine learning techniques employed to detect bots on social media platforms. The review highlights the effectiveness of algorithms such as Random Forest, Support Vector Machines (SVM), and Neural Networks, emphasizing the critical role of feature selection and data preprocessing in enhancing detection accuracy. One of the key merits of this work is its comprehensive nature, which synthesizes findings from numerous studies to present a clear overview of current methodologies and their performance metrics. However, the review also has no- table limitations, including a lack of focus on emerging challenges like adversarial bots that adapt to evade detection and insufficient exploration of practical deployment issues, such as computational costs and real-time detection capabilities in dynamic social media environments.[2]

In their Paper ”Detection and Verification of Cloned Profiles in Online Social Networks Using MapReduce-Based Clustering and Classification,” A. Saravanan and V. Venugopal (2023) present a novel methodology that employs MapReduce for the efficient detection and verification of cloned profiles on social media platforms. Their approach combines unsupervised clustering to identify similar user profiles and classification techniques to verify potential clones, allowing for scalable processing of large datasets. The primary merit of this study lies in its ability to handle big data efficiently, enhancing detection accuracy through a dual-step process that reduces false positives. the methodology also has some limitations, including its dependence on user profile attributes, which can be manipulated by cloned accounts. while the MapReduce framework enhances scalability, it may introduce latency issues in real-time applications where immediate response is critical, and the study lacks a comparative analysis of various classification algorithms that could further validate its effectiveness.[3]

In their Paper[4],(2023) investigate the various tactics employed to generate fake re- views on online platforms, providing valuable insights into the motivations and strategies of individuals and organizations behind such deceptive practices. The authors utilize a qualitative approach to categorize and analyze the production methods of fake reviews, revealing distinct strategies that range from the use of automated bots to sophisticated human manipulation. The merit of this research lies in its in-depth exploration of the complexities Title Suppressed Due to Excessive Length surrounding fake review production, highlighting the social and economic implications for businesses and consumers alike. a no- table limitation of the study is its reliance on qualitative methods, which may not capture the full extent or scale of fake review activities. while it effectively identifies strategies, the paper could benefit from quantitative analysis to measure the prevalence of these tactics and their impact on consumer behavior and business performance.[4]

In their Paper ”KC-GCN: A Semi-supervised Detection Model Against Various Group Shilling Attacks in Recommender Systems,” H. Cai, J. Ren, J. Zhao, S. Yuan, and J. Meng (2023) propose a novel detection model utilizing a knowledge-enhanced graph convolutional network (GCN) to address the challenges posed by group shilling attacks in recommender systems. The study emphasizes a semi-supervised learning approach that effectively integrates both labeled and unlabeled data, enhancing the model’s robustness in identifying fraudulent user behaviors. One of the key merits of this research is its innovative use of GCN, which leverages relational data to improve detection accuracy and adaptivity against diverse shilling attack strategies. a notable limitation is the potential reliance on the availability of quality labeled data, which can hinder performance in scenarios with limited labeled instances. Additionally, while the model shows promise in addressing group shilling attacks, further evaluation against a wider variety of attack types and in real-world applications would strengthen its applicability and generalizability.[5]

In their Paper ”Fake Profile Identification in Social Network Using Machine Learning and NLP,” Latha P and Sumitra V (2022) present a framework that leverages machine learning and natural language processing (NLP) techniques to identify fake profiles on social media platforms. The authors propose a systematic approach that combines various classification algorithms with NLP to analyze user-generated content and profile attributes, enabling the detection of deceptive accounts based on behavioral patterns and linguistic features. One significant merit of this study is its integration of NLP, which enhances the identification process by providing deeper insights into the textual characteristics of profiles, thus improving detection accuracy. However, the paper also has limitations, such as the potential challenge of generalizing the model across different social media platforms, as user behavior and content vary significantly between networks. while the study addresses the problem of fake profiles, it may benefit from a more comprehensive analysis of the underlying motivations for creating such profiles, which could inform more targeted detection strategies.[6]

In their Paper, [7] (2022) explore the application of machine learning techniques to detect and identify fake profiles on social media platforms. The authors implement a variety of classification algorithms, such as decision trees and support vector machines, to analyze user data and discern patterns indicative of fraudulent accounts. One of the key merits of this study is its practical approach, which provides a systematic evaluation of different machine learning models, showcasing their effectiveness in improving detection rates and reducing false positives in real-world scenarios. However, a notable limitation is the lack of a comprehensive dataset that represents the diversity of user behaviors across various social media platforms,

which may affect the generalizability of the model. the paper could benefit from a more detailed discussion on the ethical implications of fake profile detection, particularly concerning user privacy and data security, as well as the potential for unintended consequences in the implementation of these machine learning models.

In their Paper ”Identification of Fake Accounts in Social Media Using Machine Learning,” Kotra Shreya and Amith Kothapelly (2022) present a machine learning framework aimed at detecting fake accounts on social media platforms by analyzing user attributes and behavior patterns. The authors evaluate several classification techniques, including logistic regression and random forests, to assess their effectiveness in accurately identifying fraudulent accounts. One significant merit of this study is its thorough examination of various algorithms, providing insights into their performance metrics and applicability in real-world scenarios, which can guide future research in the field. The use of feature engineering to enhance the model’s accuracy demonstrates the authors’ commitment to a robust detection methodology. However, a limitation of the research is its reliance on a potentially narrow dataset, which may not encompass the diverse range of user behaviors across different platforms, thus affecting the model’s generalizability. the study could have explored the implications of false positives in the detection process, as misclassifying legitimate users could lead to privacy concerns and damage user trust in the platform. [8]

In their Paper [9]”Collaborative Filtering Recommendation Using Fusing Criteria Against Shilling Attacks,” L. Li, Z. Wang, C. Li, L. Chen, and Y. Wang (2022) propose a novel approach to enhance the robustness of collaborative filtering recommendation systems against shilling attacks by integrating multiple criteria for fusion. The authors present a comprehensive framework that analyzes user behavior and feedback to effectively identify and mitigate the influence of malicious users aiming to manipulate recommendations. A significant merit of this study is its innovative fusion strategy, which allows for a more accurate detection of shilling attacks, thereby improving the overall quality of recommendations and user trust. the model’s adaptability to various types of shilling attacks demonstrates its practical relevance in real-world applications. However, the pa- per also has limitations, including the complexity of the proposed method, which may lead to higher computational costs and slower response times in real-time systems. the study primarily focuses on shilling attacks without sufficiently addressing other forms of adversarial manipulations, which could limit the generalizability of the findings to broader security concerns in recommendation systems.

In their Paper ”Fake Profile Identification in Social Network Using Machine Title Suppressed Due to Excessive Length Learning and NLP,” V. Sasikala, J. Arunarasi, A.R. Rajini, and N. Nithiya (2022) explore an integrated approach combining machine learning and natural language processing (NLP) techniques to detect fake profiles on social media platforms. The authors employ various classification algorithms alongside NLP to analyze user-generated content and profile attributes, aiming to improve the accuracy of fake profile detection. One of the significant merits of this study is its dual approach, which enhances detection capabilities by leveraging both quantitative data analysis and qualitative linguistic features, thereby providing a more comprehensive assessment of user authenticity. a notable limitation is the potential for overfitting, especially if the model is trained on a limited dataset that does not adequately represent the diversity of user behavior across different social media networks. while the paper presents promising results, it could benefit from a more extensive evaluation of its methodology in real-world scenarios, considering the dynamic nature of social media and the evolving tactics used by those creating fake profiles.[10]

In their Paper[11] Shehzad et al. explores the application of big data analytics in identifying fake profiles on social media platforms. With the ever-growing number of users and data generated on social networks, fake profiles contribute significantly to the dissemination of misinformation and security risks. The authors propose a comprehensive framework leveraging big data methodologies to detect these profiles more accurately and efficiently. They utilize data from social media activities, such as behavioral patterns, account metadata, and user interactions, which are analyzed to identify

discrepancies typically associated with fake accounts. Their model incorporates machine learning algo- rithms, such as clustering and classification techniques, to categorize profiles and deter- mine anomalies indicative of inauthentic accounts. The framework is built to handle large datasets, underscoring the importance of scalable big data solutions. Results indicate that using a big data approach significantly enhances detection rates compared to traditional methods, providing an effective tool for platform administrators to monitor and mitigate risks associated with fake profiles. The study concludes that big data analytics plays a crucial role in ensuring social media security and authenticity, especially as the volume of social media data grows.

In their Paper[12]Shafiei and Dadlani address the challenge of identifying ”fickle trolls” within large-scale social networks—users who frequently change their online behavior to manipulate or disrupt conversations. Unlike traditional trolls, fickle trolls are more difficult to identify due to their adaptive strategies, making them a persistent threat to online communities. This study introduces a machine learning-based framework that combines behavioral analysis and social network data to detect these adaptive users. The model utilizes supervised learning algorithms with features derived from user activity patterns, interaction histories, and sentiment analysis to establish patterns of inconsistency. Additionally, the system is designed to function efficiently within big data environments, allowing it to operate on extensive networks without compromising accuracy. Experimental results reveal the framework’s effectiveness in identifying fickle trolls and highlight its potential for maintaining healthier online communities. This research underscores the importance of advanced analytics in tackling evolving cyber threats and emphasizes the role of machine learning in enhancing social media security.

In their Paper[13]Raza and Ding focus on the detection of fake news using a transformer- based deep learning approach. Recognizing that both the content of the news and the context in which it is shared play significant roles in spreading misinformation, the authors employ a hybrid model that analyzes both textual content and user engagement metrics. The transformer model, known for its efficiency in handling sequential data, is fine-tuned to evaluate the language patterns typical of fake news, while also considering social con- text features such as the number of shares, likes, and comments. This dual-focus approach enhances the model’s ability to distinguish between authentic and fake news with high accuracy. The study demonstrates that combining content-based and context-based features significantly improves detection outcomes compared to models that rely solely on text analysis. Raza and Ding conclude that incorporating both semantic and social cues can lead to more robust fake news detection frameworks, making it a promising solution for combating misinformation on digital platforms.

In their Paper[14]Kaushik and colleagues develop a machine learning framework to identify fake profiles on Instagram, specifically targeting platform-specific indicators of fake accounts. Instagram, as a visual-centric platform, presents unique challenges in fake profile detection, as user behaviors often differ from text-heavy platforms. The authors propose a model that integrates image analysis, user engagement metrics, and profile characteristics such as follower-following ratios and posting frequency. Using classification algorithms, the framework processes these features to identify patterns consistent with fake profiles. Their model achieved high accuracy in experimental tests, outperforming traditional detection techniques that rely on limited data points. This research highlights the potential of tailored machine learning approaches for detecting fake profiles on visually oriented social media, and it underscores the need for platform-specific adaptations in fake profile detection methods.

In their Paper [15] Vyawahare and Govilkar propose an innovative approach to fake profile recognition, focusing on linguistic features, such as profanity usage and gender identification, as key indicators. The study posits that the frequency and context of profanity, combined with inconsistencies in gender presentation, can serve as reliable indicators of fake profiles. By implementing natural language processing (NLP) techniques to analyze user posts, the authors develop a classifier that flags profiles exhibiting these traits. The framework was validated with a dataset of social network profiles, and findings indicate a significant correlation between the targeted linguistic features and inauthentic accounts. This research contributes to the growing field of fake profile detection by in- troducing unconventional yet

effective linguistic markers, suggesting that incorporating content-specific features like profanity can improve fake account detection accuracy.

## 5. METHODOLOGY

Gather a large dataset of social media profiles, including both genuine and fake accounts. This can be done by scraping social media platforms or using publicly available datasets.Clean the data by removing duplicate profiles, irrelevant information, and missing values. Convert textual data into numerical format using techniques like bag-of-words or word embeddings.Extract relevant features from the preprocessed data, such as the number of followers, engagement rate, frequency of posts, use of emojis, and language patterns.Choose a suitable machine learning algorithm for identifying fake social media accounts based on the nature of the problem. Popular algorithms include Support vector machines(SVM),K-Nearest Neighbors Algorithm(KNN),Random forest,Logistic Regression & Artificial Neural Network(ANN).Train the selected model on the preprocessed dataset using a suitable optimization technique like gradient descent or stochastic gradient descent. Split the dataset into training and testing sets to evaluate the model's performance.Evaluate the trained model's accuracy, precision, recall, and F1 score on the testing set to determine its effectiveness in identifying fake social media accounts. Use techniques like cross-validation and grid search to optimize the model's hyperparameters for better performance.Deploy the trained model in a production environment to identify fake social media accounts in real-time.Monitor its performance regularly and fine-tune it as needed to improve its accuracy over time.

## 6. PROPOSED SYSTEM ARCHITECTURE

This proposed system aims to foster better communication between the Deaf community and the hearing population by leveraging modern technologies like Deep Learning and MediaPipe, ultimately improving their social inclusion and access to information. Collaboration with stakeholders like experts in ISL, Deaf community members, and software engineers will be essential in developing an effective system.The proposed system captures live video of a user performing Indian Sign Language gestures and interprets these gestures into text. Subsequently, this text is converted into speech, allowing deaf and dumb individuals to communicate effectively. This innovation leverages deep learning and MediaPipe's advanced hand tracking capabilities to bridge communication gaps for deaf and dumb individuals performing Indian Sign Language. By providing a seamless and intuitive interface, the system could greatly enhance accessibility and communication for this community.

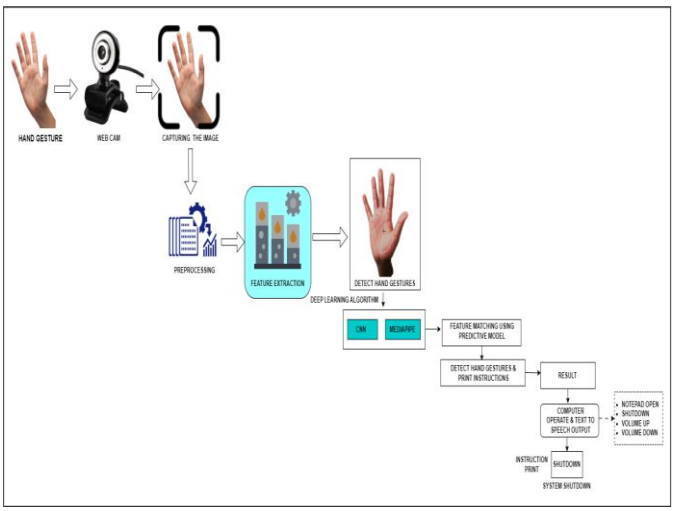


**Fig -1**: Proposed System Architecture

## 7. EXPECTED OUTCOME

The system will be able to accurately classify accounts as real or fake with high precision and recall, minimizing both false positives (flagging real accounts as fake) and false negatives (missing actual fake accounts).

By effectively identifying fake profiles, the system can contribute to reducing harmful activities such as misinformation, spamming, and scams, thereby enhancing the overall security and credibility of social media platforms.

The project will yield a comparative analysis of SVM, KNN, Random Forest, Logistic Regression, and ANN, highlighting each algorithm’s strengths and limitations in fake account detection, and informing optimal model selection or ensemble approaches.

The system will be designed to scale across social media platforms with minimal modifications, ensuring adaptability to different platform characteristics and evolving tactics used by fake accounts.

Analysis of the model's feature importance and decision-making processes will provide insights into common traits and behaviors of fake accounts, helping social media platforms understand and preemptively address emerging threats.

## 8. APPLICATION

Cybersecurity: Machine learning algorithms can be used to identify fake social media accounts that may be part of a cyber attack or phishing scam. This can help organizations to prevent data breaches and protect their users’ sensitive information.

Brand Protection: Fake social media accounts can harm a company’s reputation and damage its brand. By using

machine learning algorithms to identify fake ac- counts, companies can take proactive measures to protect their brand and prevent misinformation from spreading.

Elections and Political Campaigns: Fake social media accounts have been used to spread misinformation and manipulate public opinion during elections and polit- ical campaigns. By using machine learning algorithms to identify fake accounts, organizations can prevent the spread of false information and ensure that accurate information is presented to the public.

Fraud Detection: Fake social media accounts can be used to commit fraud, such as phishing scams or identity theft. By using machine learning algorithms to identify fake accounts, organizations can prevent fraud and protect their users' financial information.

Research and Academia: Machine learning algorithms can be used to analyze large volumes of social media data and identify patterns that may be indicative of fake accounts. This can help researchers and academics to better understand the nature of fake social media accounts and develop more effective strategies for identifying them.

## 9. CLASS DIAGRAM

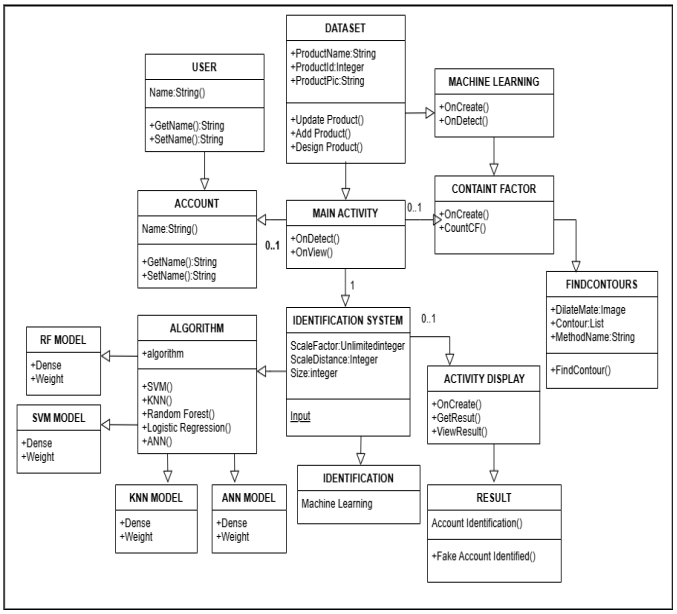


**Fig** : Class Diagram

## 10. CONCLUSION

In conclusion, the identification of fake accounts on social media using machine learn- ing algorithms such as Support Vector Machines (SVM), K-Nearest Neighbors (KNN), Random Forest, Logistic Regression, and Artificial Neural Networks (ANN) presents a promising approach to enhance the integrity and reliability of online platforms. Each algorithm offers distinct advantages in terms of accuracy, computational efficiency, and scalability, enabling the effective detection of fraudulent profiles. By leveraging a com- bination of these techniques, it is possible to develop robust systems that can adapt to the evolving nature of fake account strategies.Ongoing research is essential to refine these methods, improve their resilience against sophisticated spoofing techniques, and ensure their applicability across diverse social media environments.The successful implementa- tion of these machine learning models can contribute significantly to creating safer and more trustworthy online communities.

## ACKNOWLEDGEMENT

The development of a system architecture for identifying fake accounts on social media using machine learning draws on various interdisciplinary insights, including data science, artificial intelligence, and cybersecurity. This proposed design benefits from foundational research in machine learning techniques, specifically in natural language processing, anomaly detection, and behavior analysis, which allow for accurate profiling and detection of fake accounts based on both structured and unstructured data. Additionally, ethical considerations in data collection, feature engineering, and user privacy protection are integral to the architecture's responsible implementation, ensuring that the system not only detects fake accounts effectively but also respects platform policies and user rights.

## REFERENCES

1. Aksar, A. Firdaus, and S. A. Pasha, "Virtual vs. real self: Gendered presentation and everyday performance of virtual selfhood—A case study of Pakistan," J. Commun. Inquiry, vol. 47, no. 1, pp. 84–114, Jan. 2023, doi: 10.1177/01968599221089236.

2. M. Aljabri, R. Zagrouba, A. Shaahid, F. Alnasser, A. Saleh, and D. M. Alomari, "Ma- chine learning-based social media bot detection: A comprehensive literature review," Social Netw. Anal. Mining, vol. 13, no. 1, p. 20, Jan. 2023, doi: 10.1007/s13278-022- 01020-5.

3. A. Saravanan and V. Venugopal, "Detection and verification of cloned profiles in online social networks using MapReduce based clustering and classification," Int. J. Intell. Syst. Appl. Eng., vol. 11, no. 1, pp. 195–207, Jan. 2023.

4. S. Banerjee and A. Y. K. Chua, "Understanding online fake review pro- duction strategies," J. Bus. Res., vol. 156, Feb. 2023, Art. no. 113534, doi:10.1016/j.jbusres.2022.113534. 10 Rohini Ashok Gamane and Prof.V D dabhade
5. H. Cai, J. Ren, J. Zhao, S. Yuan, and J. Meng, "KC-GCN: A semisupervised detection model against various group shilling attacks in recommender systems," Wireless Com- mun. Mobile Comput., vol. 2023, pp. 1–15, Feb. 2023, doi: 10.1155/2023/2854874
6. Latha P, Sumitra V,"Fake Profile Identification in Social Network using Machine Learning and NLP", 2022 International Conference on Communication, Comput- ing and Internet of Things (IC3IoT)—978-1-6654-7995-0/22/31.002022IEEE—DOI : 10.1109/IC3IOT53935.2022.9767958, 978 1 6654 7995 0/22/31.00 ©2022 IEEE
7. T.Sudhakar,Bhuvana Chendrica Gogineni," FAKE PROFILE IDENTIFICATION USING MACHINE LEARNING", 2022 IEEE International Women in Engineering (WIE) Conference on Electrical and Computer Engineering (WIECON-ECE),979- 8-3503-1156-3/22/$31.00 c2022 IEEE
8. Kotra Shreya, Amith Kothapelly," Identification of Fake accounts in social media us- ing machine learning", 2022 Fourth International Conference on Emerging Research in Electronics, Computer Science and Technology (ICERECT), 978-1- 6654-5635-7/22/$31.00 ©2022 IEEE
9. L. Li, Z. Wang, C. Li, L. Chen, and Y. Wang, "Collaborative filtering recommen- dation using fusing criteria against shilling attacks," Connection Sci., vol. 34, p. 1, 1678-1696, 2022, doi: 10.1080/09540091.2022.2078280.
10. V. Sasikala, J. Arunarasi, A. R. Rajini, and N. Nithiya, "Fake profile identifica- tion in social network using machine learning and NLP," in Proc. Int. Conf. Com- mun., Comput., Internet Things (IC3IoT), Chennai, India, 2022, pp. 1–4, doi: 10.1109/IC3IOT53935.2022.9767958.
11. H. M. F. Shehzad, A. Yasin, Z. K. Ansari, M. A. Khan, and M. J. Awan, "Fake profile recognition using big data analytics in social media platforms," Int. J. Comput. Appl.Technol., vol. 68, no. 3, p. 215, 2022, doi: 10.1504/IJCAT.2022.10049746
12. H. Shafiei and A. Dadlani, "Detection of fickle trolls in large-scale online social net- works," J. Big Data, vol. 9, no. 1, p. 22, Dec. 2022, doi: 10.1186/s40537-022- 00572-9.
13. S. Raza and C. Ding, "Fake news detection based on news content and social con- texts: A transformer-based approach," Int. J. Data Sci. Analytics, vol. 13, no. 4, pp. 335–362, May 2022, doi: 10.1007/s41060-021-00302-z.
14. K. Kaushik, A. Bhardwaj, M. Kumar, S. K. Gupta, and A. Gupta, "A novelma- chine learning-based framework for detecting fake Instagram profiles,"Concurrency Comput., Pract. Exper., vol. 34, no. 28, p. e7349, Dec. 2022,doi: 10.1002/cpe.7349.
15. M. Vyawahare and S. Govilkar, "Fake profile recognition using profanity and gender identification on online social networks," Social Netw. Anal. Mining, vol. 12, no. 1, Dec. 2022, doi: 10.1007/s13278-022-00997-3.
16. B. P. Kavin, S. Karki, S. Hemalatha, D. Singh, R. Vijayalakshmi, M. Thangamani,S. L. A. Haleem, D. Jose, V. Tirth, P. R. Kshirsagar, and A. G. Adigo, "Machine learning-based secure data acquisition for fake accounts detection in future mobile communication networks," Wireless Commun. Mobile Comput., vol. 2022, pp. 1–10, Jan. 2022, doi: 10.1155/2022/6356152.