SOCIAL MEDIA FAKE ACCOUNT IDENTIFICATION USING ML

2stShraddhesh Naresh Dalvi  
*Computer Engineering Department, Sandip Institute of Engineering and Management, Nashik, India*  
dalvishraddha344@gmail.com

1stDr. Pallavi Baviskar  
*Computer Engineering Department, Sandip Institute of Engineering and Management, Nashik, India*  
[pallavi.baviskar@siem.org.in](mailto:pallavi.baviskar@siem.org.in)

3stShubham Rajendra Gholap  
*Computer Engineering Department, Sandip Institute of Engineering and Management, Nashik, India*  
[shubhamgholap2018@gmail.com](mailto:shubhamgholap2018@gmail.com)

4stPiyush Sameer Jadhav  
*Computer Engineering Department, Sandip Institute of Engineering and Management, Nashik, India*  
[piyush061103@gmail.com](mailto:piyush061103@gmail.com)

*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's 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 Networks (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.

Keywords—Support vector machines(SVM),K-Nearest Neighbors Algorithm(KNN),Random forest, Logistic Regression & Artificial Neural Network(ANN),Python

# **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 datasets representing both real and fake accounts. By training each of these algorithms on this datasets, 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 back-end 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 rising 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, salable solution, enabling platforms to respond to these sophisticated tactics. Each algorithm 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 over fitting while handling diverse features like profile metadata and user interaction patterns. Logistic Regression, known for interpret ability, 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.

# **LITERATURE REVIEW**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Sr No.** | **Title** | **Methodology** | **Findings** | **Advantages** |
| 1. | Virtual vs. real self: Gendered presentation and everyday performance of virtual self hood—A case study of Pakistan | Qualitative analysis of virtual vs. real self-presentation in a gendered context | Found distinct differences in self-presentation online, influenced by gender roles | Provides insight into how virtual identities are shaped by cultural and gender expectations, valuable for understanding fake profile motivations |
| 2. | Machine learning-based social media bot detection: A comprehensive literature review | Literature review of various ML techniques for bot detection | Summarizes existing ML approaches for detecting social media bots | Highlights effective techniques and gaps, offering a foundation for further research in social bot detection |
| 3. | Detection and verification of cloned profiles in online social networks using Map Reduce-based clustering and classification | Clustering and classification using Map Reduce | Map Reduce helps in identifying cloned profiles by grouping similar attributes | Improved efficiency for large datasets, making it suitable for big data processing |
| 4. | Understanding online fake review production strategies | Analysis of fake review generation strategies | Identifies tactics used by fake reviewers, including frequent posting and repeated content | Useful for designing fake review detection systems, improving trustworthiness of online reviews |
| 5. | KC-GCN: A semi-supervised detection model against various group shilling attacks in recommender systems | Graph Convolution Network (GCN)-based semi-supervised model | Effectively detects shilling attacks in recommendation systems | Enhances accuracy and security in recommendation algorithms, suitable for social media bot detection |

**Literature review Table 2.1**

In their Paper by Aksar, Firdaus, and Pasha (2023), titled “Virtual vs. Real Self: Gendered Presentation and Everyday Performance of Virtual Self hood 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 generalization of the findings to other societies. while the study effectively captures women’s experiences, it under develops the analysis of male virtual self hood, 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 per-processing 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 Map Reduce-Based Clustering and Classification,” A. Saravanan and V. Venugopal (2023) present a novel methodology that employs Map Reduce 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 Map Reduce 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 Recommended 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 adaptively 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 generalization. [ 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 datasets that represents the diversity of user behavior across various social media platforms, which may affect the generalization 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 datasets, 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 generalization 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 over fitting, especially if the model is trained on a limited datasets 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]

# **PROBLEM DEFINITION**

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.

# **METHODOLOGY**

1. Algorithms:

**A) Support vector machines (SVM):**

Support Vector Machines (SVMs) are a type of supervised learning algorithm that can be used for social media fake account identification using machine learning. SVMs are particularly useful in this context because they can handle high-dimensional input spaces and can provide good generalization performance. The basic idea behind SVMs is to find the best hyper plane that separates the data into two classes, such as genuine and fake accounts. The hyper plane is chosen such that the distance between it and the data points is maximized, which is known as the margin. This margin provides a measure of confidence in the classification decision. In the context of social media fake account identification, SVMs can be trained using labeled data, where each data point represents a user’s features, such as activity patterns, network structure, and content analysis. The SVM algorithm then learns to map these features onto a high-dimensional space and finds the best hyper plane that separates genuine and fake accounts. SVMs are a powerful tool for social media fake account identification using machine learning, and their ability to handle high-dimensional input spaces makes them particularly well-suited for this task.

**Algorithm 1: Working of the Support vector machines (SVM) Algorithm**

1) Collect the training data, including both features and corresponding labels indicating whether the account is fake or genuine.

2) Preprocess the data:

· Extract relevant features from user profiles, posts, interactions, etc.

· Normalize the features (e.g., using min-max scaling).

· Convert categorical features into numerical representations (e.g., one-hot encoding).

3) Split the data into training and testing sets.

4) Define the SVM model and its hyper parameters:

· Choose the type of SVM (e.g., linear, polynomial, radial basis function).

· Set the regularization parameter C, kernel parameters, and other hyper parameters.

5) Train the SVM model using the training data:

· Initialize the model with random weights and biases.

· Iterate until convergence or reaching the maximum number of iterations.

· For each training instance.

· Compute the decision function value.

· Update the model weights and biases based on misclassified instances and the regularization parameter C.

6) Evaluate the trained model using the testing set:

· For each testing instance:

· Compute the decision function value.

· Classify the instance based on the decision function value and the margin threshold.

7) Measure the performance of the model using suitable evaluation metrics (e.g., accuracy, precision, recall, F1-score) specifically tailored for fake account identification.

8) Optionally, tune the hyper parameters to improve the model’s performance:

· Perform grid search or other optimization techniques to find the best combination of hyper parameters.

1. Apply the trained model to classify new, unseen instances of social media accounts to determine if they are genuine or fake.10) Monitor and update the model periodically as new data becomes available, to maintain its effectiveness in identifying fake accounts.
2. **K-Nearest Neighbors Algorithm(KNN):**

K-Nearest Neighbors (KNN) algorithm is another machine learning algorithm that can be used for social media fake account identification. KNN is a non-parametric, lazy learning algorithm that stores all the training data and classifies new data points based on their similarity to the stored data. In social media fake account identification, KNN can be used to identify fake accounts based on their similarity to genuine accounts. The algorithm works by finding the K nearest neighbors of a new user’s features, such as activity patterns, network structure, and content analysis, and then classifying the user as genuine or fake based on the majority vote of their neighbors. One of the advantages of KNN for social media fake account identification is its ability to handle non-linear decision boundaries, as it does not assume any specific distribution of the data. However, its computational complexity can be high due to the need to store all the training data and find the nearest neighbors for each new user.KNN is a versatile algorithm for social media

fake account identification using machine learning, and its ability to handle non-linear decision boundaries makes it particularly well-suited for this task. However, its computational complexity may limit its practical use in large-scale social media platforms.

**Algorithm 2: Working of the K-Nearest Neighbors Algorithm(KNN):**

1) Data Collection: Collect a large datasets of genuine and fake social media profiles. This datasets should include features such as user demographics, profile information, activity patterns, and user interactions.

2) Prepossessing: Clean and preprocess the data to remove any missing values, duplicates, or irrelevant features. Normalize the data to ensure that all features are on a similar scale.

3) Feature Extraction: Create new features by combining existing features or applying mathematical functions to them. This step can help to improve the accuracy and efficiency of the model.

4) Model Architecture: Define the value of k (the number of nearest neighbors) and the distance metric (e.g., Euclidean distance, Manhattan distance). Use a small subset of the data for model training and validation.

5) Training: For each data point in the training set, find its k nearest neighbors based on their feature vectors using the chosen distance metric. Determine the majority class among these neighbors and assign it to the data point. Use a learning rate and batch size to control the speed and stability of the training process.

6) Testing: For each data point in the testing set, find its k nearest neighbors based on their feature vectors using the chosen distance metric. Determine the majority class among these neighbors and assign it to the data point. Evaluate the performance of the model using metrics such as accuracy, precision, recall, and F1 score. Use cross-validation to ensure that the model is robust and generalizes well to new data.

7) Deployment: Integrate the KNN into a social media platform to identify fake accounts in real-time. Monitor the performance of the model over time and retrain it periodically to adapt to new types of fake accounts or changes in user behavior.

**c) Random forest:**

Random Forest is an ensemble learning algorithm that can be used for social media fake account identification using machine learning. Random Forest is a decision tree-based algorithm that combines multiple decision trees to improve the overall accuracy and reduce over fitting. In social media fake account identification, Random Forest can be trained using labeled data, where each data point represents a user’s features, such as activity patterns, network structure, and content analysis. The Random Forest algorithm then creates multiple decision trees at training time and outputs the class that is the mode of the classes (classification) or mean of the continuous target variable (regression) for each new observation. Random Forest can also provide feature importance scores, which can help identify the most important features for social media fake account identification. This information can be used to improve the feature engineering process and reduce the dimensionality of the input space. Random Forest is a powerful tool for social media fake account identification using machine learning, and its ability to handle high-dimensional input spaces and noisy data makes it particularly well-suited for this task. However, its computational complexity may limit its practical use in large-scale social media platforms.

**Algorithm 3: Working of the Random Forest Algorithm**

1) Data Preprocessing: The first step is to preprocess the data by cleaning it, removing missing values, and normalizing the features. This step ensures that all features are on a similar scale and helps to prevent over fitting.

2) Feature Extraction: The next step is to create new features by combining existing features or applying mathematical functions to them. This step can help to improve the accuracy and efficiency of the model.

3) Model Architecture: The Random Forest algorithm consists of multiple decision trees, each trained on a different subset of the data. The number of decision trees in the forest (n estimators) is defined by the user, as well as other parameters such as the maximum depth of each tree (max depth) and the minimum number of samples required to split a node (min samples split).

4) Training: For each decision tree in the forest, a random subset of features and samples is selected from the training data. This process is called bagging (bootstrap aggregating). By training multiple trees on different subsets of the data, we can reduce over fitting and improve the robustness of the model.

5) Testing: During testing, each decision tree in the forest predicts a class for the input data point based on its feature vector. The final predicted class for the data point is determined by majority voting among all trees in the forest. This approach helps to reduce noise and improve accuracy by combining predictions from multiple trees.

6) Evaluation: The performance of the Random Forest model is evaluated using metrics such as accuracy, precision, recall, and F1 score. Cross-validation can be used to ensure that the model is robust and generalizes well to new data. The model can be retrained periodically to adapt to new types of fake accounts or changes in user behavior.

**d) Logistic Regression:**

Logistic Regression is a supervised learning algorithm that can be used for social media fake account identification using machine learning. Logistic Regression is a statistical model that predicts the probability of an event occurring based on the input features. In social media fake account identification, Logistic Regression can be trained using labeled data, where each data point represents a user’s features, such as activity patterns, network structure, and content analysis. The Logistic Regression algorithm then fits a logistic function to the data, which maps the input features to the probability of the user being genuine or fake. One of the advantages of Logistic Regression is its interpretability, as it provides coefficients for each feature that can help identify the most important factors for social media fake account identification. Logistic Regression can also handle both categorical and continuous input features, making it a versatile algorithm for social media fake account identification. However, its assumption of linearity between the input features and the output probability may limit its performance in complex non-linear relationships. Logistic Regression is a useful tool for social media fake account identification using machine learning, particularly for its interpretability and ability to handle both categorical and continuous input features. However, its assumption of linearity may limit its practical use in complex non-linear relationships.

**Algorithm 4: Working of the Logistic Regression Algorithm**

1. Data Collection: The first step is to gather a datasets that consists of various features of social media accounts that can be used to distinguish between genuine and fake accounts. Features could include the number of followers, frequency of posts, account creation date, friend count, etc. The datasets should also include a label indicating whether each account is genuine or fake.
2. Data Cleaning and Preprocessing: The datasets is then preprocessed by removing any missing values, outliers, or irrelevant features. The remaining features are scaled or normalized to ensure they have similar ranges, which helps in achieving better model performance.

3) Training and Testing Set Split: The datasets is split into a training set and a testing set. The training set is used to train or learn the model’s parameters, while the testing set is used to evaluate the model’s performance on unseen data.

4) Model Training: Logistic Regression is a supervised learning algorithm, meaning it requires labeled data to learn from. The algorithm estimates the relationship between the features and the binary outcome using a logistic function. By fitting the logistic curve to the training data, the model learns

the optimal weights (coefficients) for each feature, minimizing the error between the predicted outcome and the actual label.

5) Model Evaluation: Once the model is trained, it is evaluated using the testing set to assess how well it generalizes to new, unseen data. Common evaluation metrics for binary classification include accuracy, precision, recall, and F1 score. These metrics quantify the model’s performance in terms of correctly identifying genuine and fake accounts.

6) Hyper parameter Tuning: Hyper parameters are adjustable settings that control the learning process of the algorithm. Examples of logistic regression hyper parameters include regularization strength and learning rate. Hyper parameter tuning techniques like grid search or random search can be applied to find the optimal combination of hyper parameters that maximize the model’s performance.

Model Deployment: After the model is trained and evaluated, it can be deployed to make predictions on new, unlabeled data. Given a set of features for an unknown social media account, the model can predict the probability of it being genuine or fake based on the learned weights and logistic function.

**e) Artificial Neural Network (ANN):**

Artificial Neural Network (ANN) is a machine learning algorithm inspired by the structure and functioning of the human brain. It consists of an interconnected network of artificial neurons, called nodes or artificial neurons that work together to process and transmit information. The algorithm is composed of multiple layers an input layer, one or more hidden layers, and an output layer. Each layer contains a number of nodes that are interconnected with weighted connections. The input layer receives input data, which is then processed and transmitted through the hidden layers until it reaches the output layer, where the final prediction or decision is made. ANNs are commonly used for tasks such as classification, regression, clustering, and pattern recognition. They can learn from large datasets and generalize pat-terns, making them useful in a wide range of applications including image and speech recognition, natural language processing, and predictive modeling. ANNs have the ability to learn complex patterns and can adapt to new data, but they can also be prone to over fitting and require careful tuning of hyper parameters.

**Algorithm 5: Working of the Artificial Neural Network (ANN) Algorithm**

1) Define the neural network architecture:

o Input layer: number of features in the input data

o Hidden layer(s): number of hidden layers and number of neurons in each layer

o Output layer: single neuron representing the probability of being a fake account

2) Initialize the weights and biases of the neural network randomly.

3) Specify the activation function for the neurons (e.g., sigmoid, ReLU, etc.).

4) Set the learning rate, number of iterations, and other hyper parameters for the neural network training.

5) Preprocess the input data:

o Normalize the numerical features (e.g., using min-max scaling)

o Convert categorical features into numerical representations (e.g., one-hot encoding)

6) Split the data into training and testing sets.

7) Train the neural network using a suitable optimization algorithm (e.g., gradient descent):

o Forward propagation:

o Compute the weighted sum of inputs and biases in each layer

o Apply the activation function to obtain the output of each neuron

8) Calculate the loss using a suitable loss function (e.g., binary cross-entropy)

Back propagation:

o Compute the gradients of the loss with respect to the weights and biases

o Update the weights and biases using the gradients and the learning rate

9) Iterate steps 7 until convergence or reaching the maximum number of iterations.

10) Evaluate the trained model using the testing set:

o Forward propagate the test data through the trained neural network

o Classify each instance as a fake or genuine account based on the output probability threshold

1. Measure the performance of the model using suitable evaluation metrics (e.g., accuracy, precision, recall, F1-score)..

# **SYSTEM ARCHITECTURE**

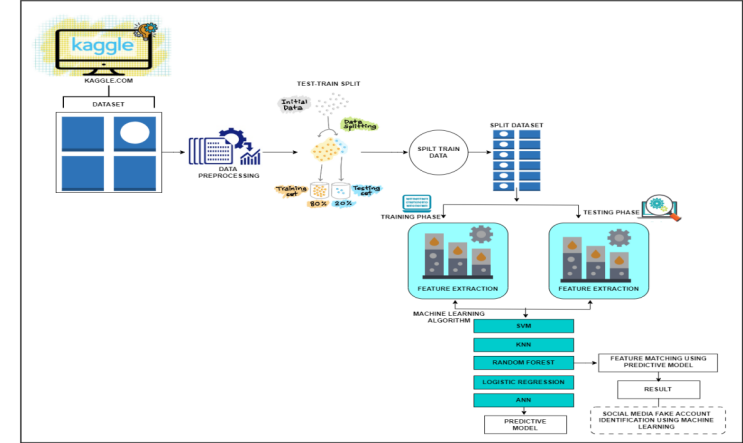
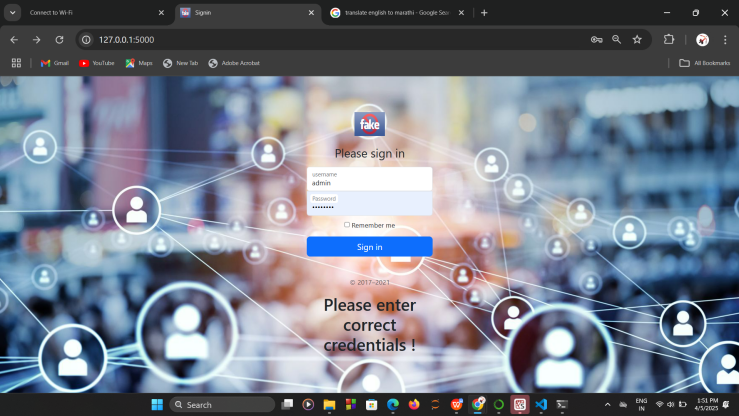


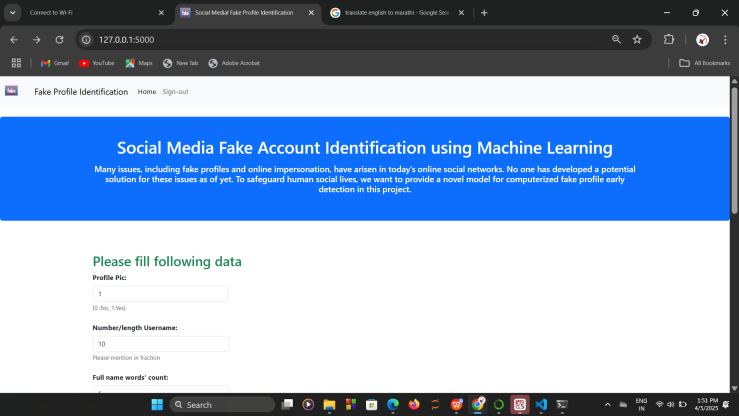
Fig 5.1

This proposed system aims to foster better communication between the Deaf community and the hearing population by leveraging modern technologies like Deep Learning and Media Pipe, 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 Media Pipe’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.

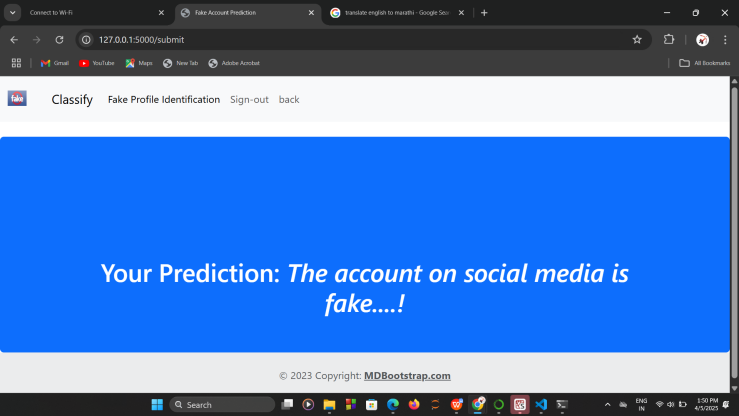
# **RESULTS**

****

Login Page fig.6.1

****

Filling Data fig.6.2



Final Output fig.6.3

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 behavior of fake accounts, helping social media platforms understand and pre-emotively address emerging threats.

# **ACKNOWLEDGMENT**

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 cyber security. 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.

# **CONCLUSION**

In conclusion, the identification of fake accounts on social media using machine learning 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 combination 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 implementation of these machine learning models can contribute significantly to creating safer and more trustworthy online communities.

# **REFERENCES**

1. Aksar, A. Firdaus, and S. A. Pasha, “Virtual vs. real self: Gendered presentation and everyday performance of virtual self hood—*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, “*Machine 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, Computing 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 using 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.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 identification 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/IC3IOT53935.2022.9767958.1.

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

1. 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.
2. A. Saravanan and V. Venugopal, “*Detection and verification of cloned profiles in online social networks using Map Reduce based clustering and classification*,” Int. J. Intell. Syst. Appl. Eng., vol. 11, no. 1, pp. 195– 207, Jan. 2023.
3. S. Banerjee and A. Y. K. Chua, “*Understanding online fake review production 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
4. 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

15.Latha P, Sumitra V,”*Fake Profile Identification in Social Network using Machine Learning and NLP*”, 2022 International Conference on Communication, Computing 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

16.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

17. Kotra Shreya, Amith Kothapelly,” *Identification of Fake accounts in social media using 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

18 .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.

19. V. Sasikala, J. Arunarasi, A. R. Rajini, and N. Nithiya, “*Fake profile identification 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. Pallavi Vasudev Baviskar " \_*Design of Machine Learning-Based Malware Detection Techniques in Smartphone Environment\_* " in 2nd International Conference for Advancement in Technology (ICONAT) during 24th & 26th January 2023.