Pooja Dharshini K, 2026, 14:2
ISSN (Online): 2348-4098
ISSN (Print): 2395-4752

International Journal of Science, Engineering and Technology
An Open Access Journal

Fake Social Media Profile Detection Using Machine Learning

Pooja Dharshini K
Computer Science and Engineering St. Peter's Institute of Higher Education and Research

Abstract- Fake profiles are often created to spread misinformation, disinformation, and propaganda. Detecting and removing these profiles is crucial to curb the dissemination of false information and maintain the integrity of information shared on social media. In this study, we present an enhanced algorithm for the detection of fake social media profiles, utilizing machine learning techniques such as Gradient Boosting, Random Forest, and Support Vector Machine. The algorithm incorporates a range of profile features, including the presence of a profile picture, characteristics of the username and full name (length, numbers, equality), description length, external URL presence, account privacy, and key metrics like the number of posts, followers, and follows. The primary objective is to address the escalating issue of fraudulent activities and misinformation on social media platforms. The proposed algorithm leverages ensemble learning to improve the accuracy and reliability of identifying deceptive profiles. Additionally, we introduce a Flask-based web application to deploy the Random Forest algorithm, enabling real-time detection and providing a user-friendly interface. To evaluate the algorithm's performance, precision, recall, and F1 score are employed as key metrics. Precision measures the accuracy of positive predictions, recall gauges the algorithm's ability to capture all positive instances, and the F1 score balances precision and recall. Through comprehensive testing and validation, our algorithm aims to contribute to the advancement of online security, fostering user trust and mitigating the impact of fake profiles in the dynamic landscape of social media.

Keywords: Fake social media profiles , Machine learning , Detection , Supervised learning , Support vector machine , Neutral networks , Profile metadata, logistics regression , Deep

I. INTRODUCTION OVERVIEW

Fake profiles are commonly used for fraudulent activities and scams. Cybercriminals may impersonate individuals or organizations to deceive users, leading to financial losses, identity theft, or other malicious activities. Identifying and eliminating fake profiles helps protect users from falling victim to scams. Machine learning (ML) plays a crucial role in detecting fake profiles on social media platforms by leveraging patterns and characteristics learned from labeled data. Machine learning algorithms analyze various features extracted from user profiles, posts, and activities. ML models excel at recognizing complex patterns and relationships within data. By training on labeled datasets containing examples of both genuine and fake profiles, these models learn to identify subtle and nuanced characteristics associated with deceptive accounts.

In the current generation, everyone's social life is now entwined with online social networks. It has been simpler to add new friends and stay in touch with them and their updates. Online social networks influence a variety of fields, including research, education, community activism, employment, and business. These online social networks have been the subject of research to see how they affect people. Instructors may quickly contact their students using this, creating a welcoming atmosphere for them to learn. Teachers are becoming more familiar with these sites and using them to provide online classroom pages, assign assignments, hold conversations, and other activities that greatly enhance learning. Employers may utilize these social networking sites to find and hire skilled individuals who are enthusiastic about their jobs. In addition to all of this, propaganda spread via social media is a problem. False accounts sharing incorrect and unsuitable information cause conflicts.

The rapid growth of social media has transformed how we communicate, interact, and share information. However, this digital revolution has also opened new avenues for malicious activities. One significant challenge is the rise of fake profiles on social media platforms, which can spread misinformation, engage in fraudulent activities, manipulate public opinion, and contribute to other harmful behaviors.

OBJECTIVE

The main objective of the study is to develop and deploy a robust algorithm for detecting fake profiles on social media platforms, leveraging advanced machine learning techniques like Gradient Boosting, Random Forest, and Support Vector Machine.

CHALLENGES IN THE DOMAIN

Scale and Volume: Social media platforms host billions of accounts, making manual review and monitoring impractical. The sheer volume of data necessitates automated approaches, but this also increases the complexity of analysis.

Sophistication of Fake Profiles: Fake profiles are becoming increasingly sophisticated, often mimicking genuine user behaviour. This makes it harder to detect them based on simple indicators. Malicious actors may use AI to create realistic content, deepfake images, and other techniques to avoid detection.

Dynamic and Evolving Threats: The tactics used to create and manage fake profiles are continuously evolving. Methods that work today may become obsolete as malicious actors adapt to new detection techniques, requiring constant algorithm updates.

Diverse Use Cases: Fake profiles can be used for a wide range of purposes, from spreading misinformation and propaganda to conducting fraud and identity theft. This diversity complicates the task of developing a one-size-fits-all solution.

PROBLEM STATEMENT

- **Proliferation of Fake Profiles:**

The rapid increase in fake profiles on social media is leading to widespread misinformation, fraudulent activities, and manipulation of public opinion.

- **Insufficient Traditional Detection Methods:**

Manual review and user reporting are inadequate to handle the sheer volume of accounts and the evolving tactics of fake profile creators.

- **Sophistication of Deceptive Tactics:**

Fake profiles are becoming more sophisticated, using techniques like AI-generated content, making them harder to detect with basic automated tools.

- **False Positives and False Negatives:**

Effective detection must minimize both false positives (genuine accounts incorrectly flagged as fake) and false negatives (fake profiles that go undetected).

II. LITERATURE SURVEY

T. Sudhakar, B. C. Gogineni and J. Vijaya, "Fake Profile Identification Using Machine Learning," 2022 IEEE International Women in Engineering (WIE) Conference on Electrical and Computer Engineering (WIECON-ECE), Naya Raipur, India, 2022, pp. 47-52, Doi: 10.1109/WIECON-ECE57977.2022.10150753.

Online social networks have permeated our social lives in the current generation. These sites have allowed us to see our social lives differently than they did in the past. Nowadays we can connect with new friends and maintain relationships with them via social and personal activities become quite easy. Online Social Networks (OSN) are contributed in all areas such as Research in all domains, Job-related areas, Technology oriented areas, Health care, and business-oriented areas, Information gathering and data collection, and so on. One of the biggest problems on these social media platforms is fake profiles. Impersonating to be someone else and causing harm and defamation to the real person or advertising or popularizing removed propaganda on someone's name to get more benefit is the motto of such profile creators. There have been many studies regarding these fake accounts and how can they be mitigated. Many approaches such as graph-level activities or feature analysis have been taken into consideration to identify fake profiles. These methods are outdated when compared to arising issues of these days. In this paper, we proposed a technique using machine learning for fake profile detection which is efficient. . The benchmark data set is collected and mixed with manual data first furthermore; a data cleaning technique is used to present the data more feasibly. Then the preprocessed data is used for model building with sufficient information such as profile name, profile ID name, number of followers, and so on. We added Cross validation process where many training algorithms are implemented on the given data and are then tested on the same data. Based on the experiments the RF classifier performed better than the other classification methods. The Random Forest classifier is used to forecast the profile whether is fake or genuine in an efficient way.

K. Patel, S. Agrahari and S. Srivastava, "Survey on Fake Profile Detection on Social Sites by Using Machine Learning Algorithm," 2020 8th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO), Noida, India, 2020, pp. 1236-1240, Doi: 10.1109/ICRITO48877.2020.9197935.

To avoid the spam message, malicious and cyber bullies activities which are mostly done by the fake profile. These activities challenge the privacy policies of the social network communities. These fake profiles are responsible for spread false information on social communities. To identify the fake profile, duplicate, spam and bots account there is much research work done in this area. By using a machine-learning algorithm, most of the fake accounts detected successfully. This paper represents the review of Fake Profile Detection on Social Site by Using Machine Learning.

L. P, S. V, V. Sasikala, J. Arunarasi, A. R. Rajini and N. Nithiya, "Fake Profile Identification in Social Network using Machine Learning and NLP," 2022 International Conference on Communication, Computing and Internet of Things (IC3IoT), Chennai, India, 2022, pp. 1-4, Doi: 10.1109/IC3IOT53935.2022.9767958.

In present times, social media plays a key role in every individual life. Everyday majority of the people are spending their time on social media platforms. The number of accounts in these social networking sites has dramatically increasing day-by-day and many of the users are interacting with others irrespective of their time and location. These social media sites have both pros and cons and provide security problems to us also for our information. To scrutinize, who are giving threats in these networking sites we need to organize these social networking accounts into genuine accounts and fake accounts. Traditionally, we are having different classification methods to point out the fake accounts on social media. But we must increase the accuracy rate in identifying fake accounts on these sites. In our paper we are going with Machine Learning technologies and Natural Language processing (NLP) to increase the accuracy rate of detecting the fake accounts. We opted for Random Forest tree classifier algorithm.

R. Bhambulkar, S. Choudhary and A. Pimpalkar, "Detecting Fake Profiles on Social Networks: A Systematic Investigation," 2023 IEEE International Students' Conference on Electrical, Electronics and Computer Science (SCEECS), Bhopal, India, 2023, pp. 1-6, Doi: 10.1109/SCEECS57921.2023.10063046.

Social media platforms are an excellent way to remain in touch with loved ones. Still, they can also be a breeding ground for fake accounts. These accounts may be used to spread malicious content or disinformation or to manipulate people's thoughts and feelings. One way to identify fake social media accounts is to use machine learning algorithms. Machine learning enables computers to learn without being explicitly programmed, automatically recognizing patterns in data. This paper examines several methods implemented and tested by machine learning and deep learning algorithms. We have also looked into the numerous aspects associated with user profiles on social networks.

M. S. Kumar, J. Sabeena, K. M. Veena, K. Pavan, M. Sukavya and K. Sravanthi, "Fake Profile Detection on Social Networking Websites using Machine Learning," 2023 International Conference on Sustainable Computing and Smart Systems (ICSCSS), Coimbatore, India, 2023, pp. 119-122, Doi: 10.1109/ICSCSS57650.2023.10169168.

These days, social media has a significant impact on everyone's life. Most people frequently utilize social media platforms. Each of these social media platforms offers benefits and drawbacks, as well as security risks for our information. To determine who poses threats on these platforms, it is necessary to distinguish between the real and fake social media profiles. There are traditionally used various methods for identifying fake social media accounts. But these platforms need to be better at identifying phone accounts. The accuracy rate of identifying fake accounts utilizing timestamp data types is improved in this proposed work employing high gradient boosting algorithms and Natural Language Processing. In order to investigate the relationship between various machine learning methods and multi-features in time series, this study employs a variety of machine learning techniques.

S. Khaled, N. El-Tazi and H. M. O. Mokhtar, "Detecting Fake Accounts on Social Media," 2018 IEEE International Conference on Big Data (Big Data), Seattle, WA, USA, 2018, pp. 3672-3681, Doi: 10.1109/BigData.2018.8621913.

In the present generation, on-Line social networks (OSNs) have become increasingly popular, people's social lives have become more associated with these sites. They use on-Line social networks (OSNs) to keep in touch with each other's, share news, organize events, and even run their own e-business. The rapid growth of OSNs and the massive amount of personal data of its subscribers have attracted attackers, and imposters to steal personal data, share false news, and spread malicious activities. On the other hand, researchers have started to investigateefficient techniques to detect abnormal activities and fake accounts relying on accounts features, and classification algorithms. However, some of the account's exploited features have negative contribution in the final results or have no impact, also using standalone classification algorithms does not always achieve satisfactory results. In this paper, a new algorithm, SVM-NN, is proposed to provide efficient detection for fake Twitter accounts and bots, feature selection and dimension reduction techniques were applied. Machine learning classification algorithms were used to decide the target account's identity real or fake, those algorithms were support vector machine (SVM), neural Network (NN), and our newly developed algorithm, SVM-NN. The proposed algorithm (SVM-NN) uses a smaller number of features, while still being able to correctly classify about 98% of the accounts of our training dataset.

P. K. Roy and S. Chahar, "Fake Profile Detection on Social Networking Websites: A Comprehensive Review," in IEEE Transactions on Artificial Intelligence, vol. 1, no. 3, pp. 271-285, Dec. 2020, Doi: 10.1109/TAI.2021.3064901.

Social networking platforms have become an essential part of today's human life—almost every individual is associated with at least one of the online social networking websites today. Hence, a huge crowd is always active on these platforms; a large number of user engagements attracted spammers and unauthentic users on online social networking. To spread unauthentic messages such as rumors, hate speech, bullied text, and others, users create a fake profile. Researchers proposed several techniques to limit this issue using machine-learning- and deep-learning-based models, but many fake accounts are still present. However, for a good social networking platform, these fake accounts are not acceptable. This article summarizes the recent advancement of social networking's fake account detection, which helps the future researcher build a robust model to prevent and identify fake accounts on online social networking.

In recent years, telecommunication has become the major media of communication, while the market has been saturated in many developed countries. Telecommunication companies are shifting their concerns from acquiring new customers to retaining regular customers, since the cost to win a new customer is higher than that of maintaining a regular customer. Therefore, it is essential for telecommunication operators to forecast customer churn in this competitive, accurate market. This paper aims to identify affecting customer churn and construct an efficient model, which is used to predict and analyze data with visualization results. The churn forecast consists of several phases: data prepossessing, data analysis, evaluation measure, and application of machine learning algorithms. Moreover, data pre-processing covers data cleaning, transformation, and classification. The machine learning classifiers selected are Logistic Regression, SVM, Random Forest, AdaBoost, GBDT, XGBoost, Light GBM, and CatBoost. Classifiers were evaluated using performance measures, such as accuracy, precision, recall, AUC, and F1-Score. Based on the paper, the result was shown that the Light GBM outperformed other classifiers while identifying potential churners.

N. Singh, T. Sharma, A. Thakral and T. Choudhury, "Detection of Fake Profile in Online Social Networks Using Machine Learning," 2018 International Conference on Advances in Computing and Communication Engineering (ICACCE), Paris, France, 2018, pp. 231-234, Doi: 10.1109/ICACCE.2018.8441713.

In today's world, the social media platforms are being used on daily basis and has become an important part of our lives. The number of peoples on social media platforms are incrementing at a greater level for malicious use. There are numerous cases where produced accounts have been effectively distinguished utilizing machine adapting techniques however the amount of research work is very low to recognize counterfeit characters made by people. For bots the ML models used various features to calculate the no. of followers to the no. of friends that an account has on social media platforms (SOCIAL MEDIA PLATFORMS). The no. of friends to the no. of followers of any account are easily available in the account profiles and no rights are violated of an accounts.

III. EXISTING SYSTEM

Detecting fake social media profiles using machine learning typically involves analyzing various features and patterns within the profiles to differentiate between genuine and fake accounts. Here's a high-level overview of how such a system might work:

- **Data Collection***: Gather a dataset of labeled social media profiles, with labels indicating whether each profile is genuine or fake. This dataset should encompass a wide range of features, including profile information, activity patterns, posting behavior, network connections, etc.

- **Feature Extraction***: Extract relevant features from the social media profiles. These features could include profile completeness, posting frequency, engagement patterns (likes, comments), language usage, network characteristics (number of friends/followers, network diversity), and so on.

- **Preprocessing*:** Clean and preprocess the data, which may involve handling missing values, normalizing numerical features, encoding categorical variables, and dealing with outliers.

- **Model Selection***: Choose appropriate machine learning models for classification, such as decision trees, random forests, support vector machines (SVM), logistic regression, or neural networks. Ensemble methods like boosting or bagging can also be effective.

- **Training***: Train the selected model(s) on the labeled dataset, using techniques like cross-validation to optimize hyperparameters and ensure generalization.

- **Evaluation*:** Evaluate the trained models using performance metrics such as accuracy, precision, recall, F1-score, and ROC-AUC to assess their effectiveness in distinguishing between genuine and fake profiles.

- **Deployment***: Once a satisfactory model is trained, deploy it into a production environment where it can automatically analyze and flag potentially fake profiles in real-time or in batches.

- **Monitoring and Maintenance*:** Continuously monitor the performance of the system in detecting fake profiles and periodically retrain the model with new data to adapt to evolving tactics used by malicious actors. It's worth noting that detecting fake profiles is a challenging and evolving task, as malicious actors constantly adapt their strategies to evade detection. Therefore, a robust system may need to incorporate additional techniques such as natural language processing (NLP), anomaly detection, and network analysis to stay ahead of emerging threats.

Proposed Sytem

The proposed system for predicting and minimizing customer churn in banks using machine learning involves several key steps.

➢ **Data Collection:**
- Gather a diverse and representative dataset of social media profiles, including both genuine and fake profiles.

➢ **Feature Engineering:**
- Identify and extract relevant features from social media profiles that can be used to distinguish between genuine and fake accounts.
- Features may include profile picture presence, username and full name characteristics, description length, external URL presence, account privacy, and key metrics like the number of posts, followers, and follows.

➢ **Data Preprocessing:**
- Handle missing or inconsistent data.
- Normalize or scale numerical features.
- Encode categorical features.
- Split the dataset into training and testing sets.

➢ **Algorithm Selection:**
- Choose machine learning algorithms suitable for the task. In this case, Gradient Boosting, Random Forest, and Support Vector Machine are mentioned.

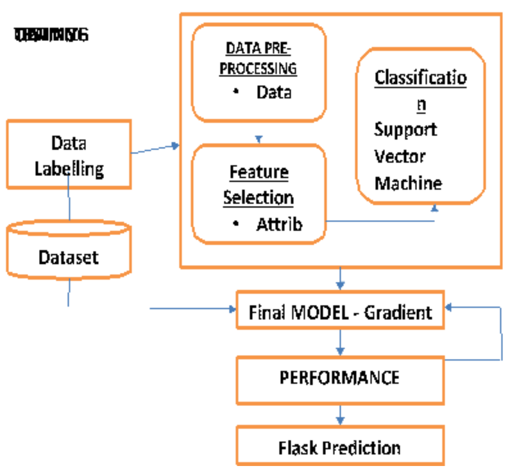
➢ **Model Training:**
- Train the selected machine learning models using the training dataset.

➢ **Model Evaluation:**
- Assess the performance of the trained models using metrics such as precision, recall, and F1 score.
- Compare the results of different algorithms and ensemble methods.

➢ **Web Application Development:**
- Implement a Flask-based web application to deploy the Random Forest algorithm.
- Create a user-friendly interface for real-time detection of fake profiles.

Modules

- Data collection
- Data analysis
- Data preprocessing
- Model training
- Model testing/performance evaluation



Data Collection

- The data set includes information about: Target Variable (fake): Indicates whether a profile is fake (1) or not fake (0).

- Features: These can include a mix of numerical and categorical data points, such as:
- Profile Picture: Whether the profile has a profile picture (binary, 1 for Yes, 0 for No).
- Numbers in Username: The ratio of numbers to the total length of the username.
- Words in Full name : Number of words in the user's full name.
- Numbers in Full name : The ratio of numbers to the total length of the full name.
- Username Equals Name: Whether the username is the same as the user's full name (binary).
- Description Length: The length of the user's profile description.
- External URL: Whether the profile has an external URL (binary).
- Private Profile: Whether the profile is private (binary).

- Number of Posts: The total number of posts made by the profile.
- Number of Followers: The number of followers the profile has.
- Number of Follows: The number of other profiles this profile follows.

Data Preprocessing

- Loading data from a CSV file.
- Checking for null values and basic data types.
- Addressing outliers with log transformation.
- Splitting data into training and testing sets.
- Standardizing features to ensure consistent scaling.

Data Loading
Reading CSV Data: Load the data from a CSV file into a Pandas Data Frame. This is typically the first step in data preprocessing.

Data Inspection
Checking for Null Values: Identify any missing data in the Data Frame.

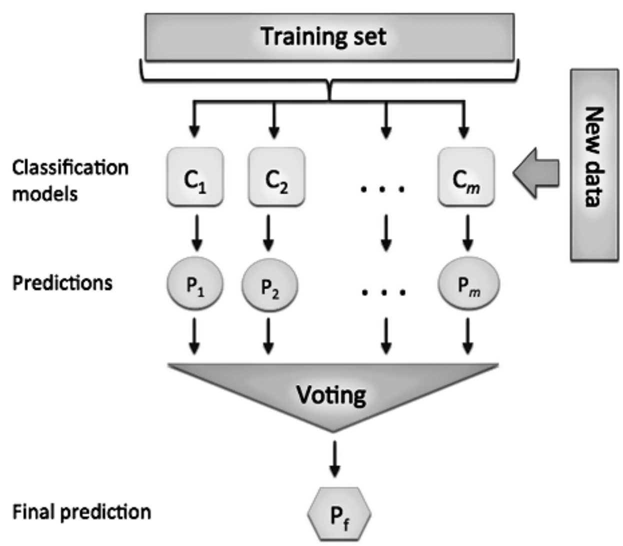
Outlier Detection and Handling
Log Transformation: To address outliers, you applied a logarithmic transformation to specific features that are prone to having extreme values. This reduces the skewness of the distribution.

Train-Test Split
Splitting Data into Training and Testing Sets: Separate the dataset into training and testing subsets for model training and evaluation.

Data Standardization
Standardizing the Data: Standardization ensures that the features have a mean of 0 and a standard deviation of 1. This step is particularly important for certain machine learning algorithm

MODEL TRAINING

- Gradient Boosting: A boosting technique that builds trees sequentially to improve model accuracy.
- Random Forest: An ensemble method that constructs multiple trees in parallel and combines their outputs for prediction.
- Support Vector Machine: A method that finds an optimal hyperplane for classification, with flexibility through kernel functions.



- **Gradient Boosting Classifier**

Algorithm: Gradient Boosting is an ensemble learning method that builds multiple decision trees sequentially, with each new tree attempting to correct the errors made by the previous ones. This approach often leads to high accuracy and flexibility.

Advantages: Typically achieves high accuracy, robust to overfitting if tuned properly, and allows for feature importance analysis.

- **Random Forest Classifier**

Algorithm: Random Forest is also an ensemble learning method, but it builds multiple decision trees in parallel and combines their outputs to make predictions. It introduces randomness to improve robustness and reduce overfitting.

Advantages: Generally robust to overfitting, provides feature importance, and handles high-dimensional data well.

- **Support Vector Machine (SVM)**

Algorithm: SVM finds a hyperplane that best separates different classes in a high-dimensional space. It uses kernel methods to handle non-linear relationships.

Advantages: Effective in high-dimensional spaces, versatile with different kernels, and often achieves good performance.

IV. UML DIAGRAM

Artificial Intelligence

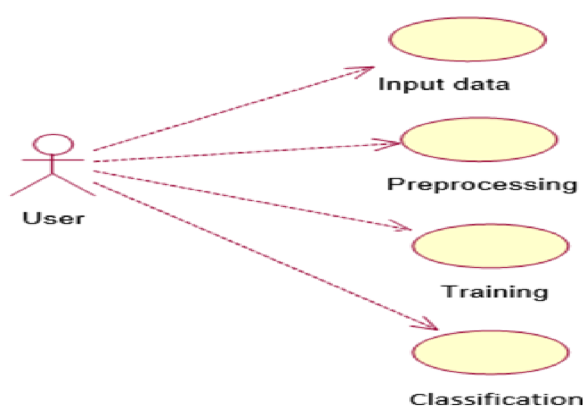
Artificial intelligence (AI) is the simulation of human intelligence processes by machines, especially computer systems enabling it to even mimic human behavior. Its applications lie in fields of Computer Vision, Natural Language Processing, Robotics, Speech Recognition, etc. Advantages of using AI are improved customer experience, accelerate speed to market, develop sophisticated products, enable cost optimization, enhance employee productivity and improve operational efficiency. Machine Learning (ML) is a subset of AI which is programmed to think on its own, perform social interaction, learn new information from the provided data and adapt as well as improve with experience. Although training time via Deep Learning (DL) methods is more than Machine Learning methods, it is compensated by higher accuracy in the former case. Also, DL being automatic, large domain knowledge is not required for obtaining desired results unlike in ML.

Uml Diagrams:

Unified Modelling Language (UML) is simply another graphical representation of a common semantic model. The proposed system has been designed by using use case diagram, class diagram, sequence diagram, collaboration diagram, state chart diagram and component diagram.

Use Case Diagram

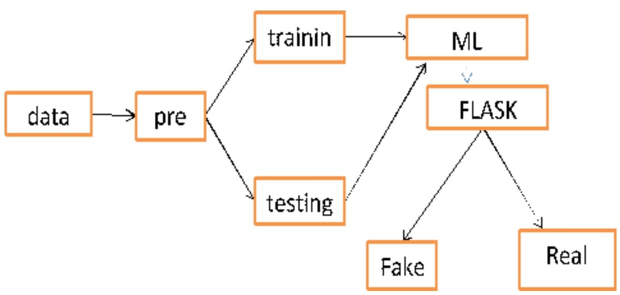
The use case diagram consists of the actors and the use cases. The actors of the system are user, system holder, device controller and the use cases are authentication, checking credentials, basic ON/OFF, allow/deny user, storing NLP commands, Input through voice commands, Deriving Data, Intrusion Detection, Service Maintenance. describes Use Case diagram for Adaptive Automation System (AAS).User and Non-User can login to the device by Admin. Then device controller decides the entry of the user. The user can control the devices through voice commands and these can be monitored by Device Controller. Intrusion detection and maintenance also controlled by the Device Controller (Admin). The user can check for the intruder in the environment and based on the intruder the door can be open/closed. The light can be turned on/off by the system holder and they can alert the emergency system if any poisonous gas or temperature exceeds the normal level. This can be maintained by system maintenance.



Class Diagram

Class diagram is to model the static view of an application. Class diagrams are the only diagrams which can be directly mapped with object-oriented languages and thus widely used at the time of construction and it is used for general conceptual modelling of the structure of the application, and for detailed modelling translating the models into programming code.

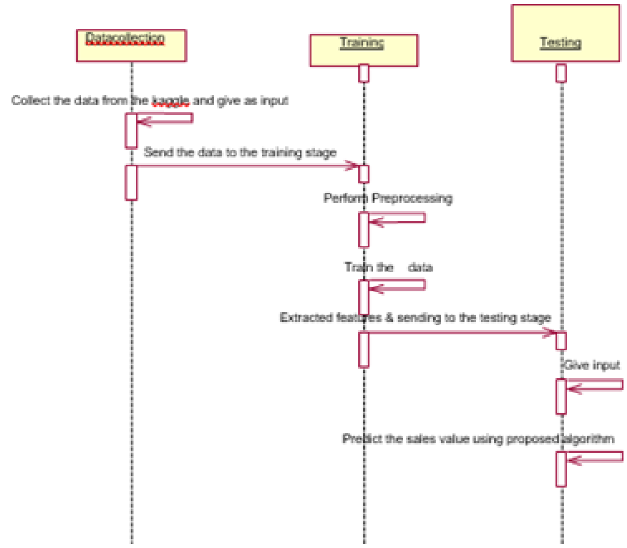
Class diagrams can also be used for data modelling. The classes in a class diagram represent both the main elements, interactions in the application, and the classes to be programmed.



Sequence Diagram

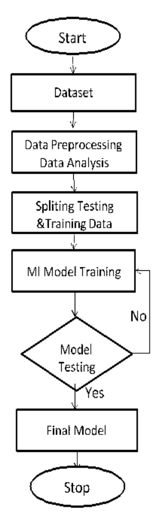
The control flow between various participants or entity roles of the corresponding system in the form of messages is represented in the Sequence Diagram. The participants are represented within the rectangular object. The swim line or the lifeline that is dragged below every participant represents the lifetime of the corresponding participant.

The UML representation of a class is rectangle containing three compartments stacked vertically. The top compartments show the class's name. the middle compartments list the class's attributes. The bottom compartment lists the class operations known as the methods of the class. A class diagram consists of any number of classes which will be connected by the lines, which may have arrows at one or both ends, connecting the boxes. These lines define the relationships, also called associations, between the classes. These lines will have multiplicity to represent the number of instances of the classes.



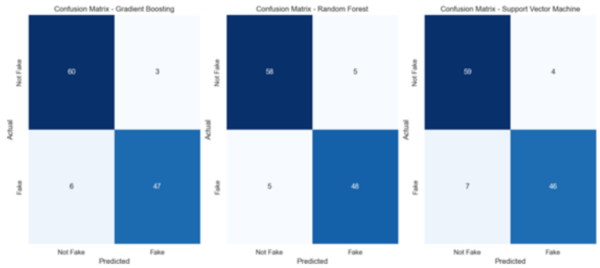
Data Flow Diagram:

A Data Flow Diagram (DFD) is a traditional way to visualize the information flows within a system. A neat and clear DFD can depict a good amount of the system requirements graphically. It can be manual, automated, or a combination of both. It shows how information enters and leaves the system, what changes the information and where information is stored. The purpose of a DFD is to show the scope and boundaries of a system as a whole. It may be used as a communications tool between a systems analyst and any person who plays a part in the system that acts as the starting point for redesigning a system.



V. OUTPUT SCREENSHOT

General



- **Accuracy**

Definition: Accuracy is the ratio of correct predictions (both true positives and true negatives) to the total number of predictions. It indicates how often the model correctly predicts the correct class.
Interpretation: Higher accuracy suggests that the model is generally performing well. However, accuracy can be misleading if there is a class imbalance, as it doesn't distinguish between correct identification of positive and negative classes.
Precision

Definition: Precision measures the proportion of true positive predictions out of all predicted positives. It tells us how many of the detected positive results were actually correct.
Interpretation: A higher precision means fewer false positives (incorrectly classified positive results). In the context of fake profile detection, high precision indicates that when the model identifies a profile as fake, it is likely to be correct.

- **Recall**

Definition: Recall, also known as sensitivity or true positive rate, measures the proportion of true positive results out of all actual positive instances. It reflects the model's ability to identify all relevant instances of the positive class.
Interpretation: High recall suggests that the model effectively identifies fake profiles, with fewer false negatives (missing actual fake profiles). A high recall is crucial in applications where missing a positive instance could lead to significant risks or consequences.

- **F1 Score**

Definition: The F1 score is the harmonic mean of precision and recall, providing a balanced measure. It is used to evaluate the model's overall performance, especially when precision and recall need to be balanced.
Interpretation: A higher F1 score indicates a good balance between precision and recall. It's particularly useful when the dataset has a class imbalance, as it accounts for both false positives and false negatives.
Applying the Metrics to the Models

- **Gradient Boosting:**

Shows high accuracy (0.922414), indicating it performs well overall.
High precision (0.94000) suggests that when it predicts a profile as fake, it's likely to be correct. Strong recall (0.886792) implies it can capture most of the actual fake profiles.

The F1 score (0.912621) reflects a good balance between precision and recall, indicating the model is robust.

- **Random Forest:**

Slightly lower accuracy (0.913793) than Gradient Boosting but still high.

Precision (0.90566) and recall (0.905660) are closely matched, suggesting it maintains a balance between avoiding false positives and capturing positive instances.

F1 score (0.905660) aligns with its balanced precision and recall, indicating consistent performance.
Support Vector Machine (SVM):

Accuracy (0.905172) is slightly lower than the other two models but still high.

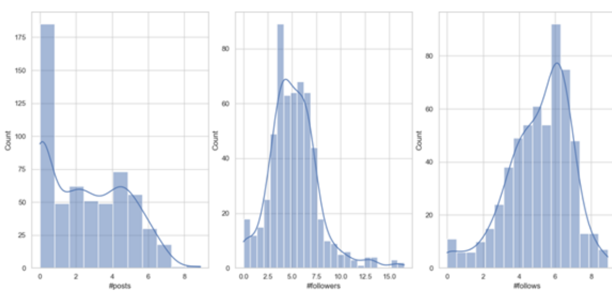
Precision (0.92000) is quite high, indicating it tends to be accurate in its positive predictions.

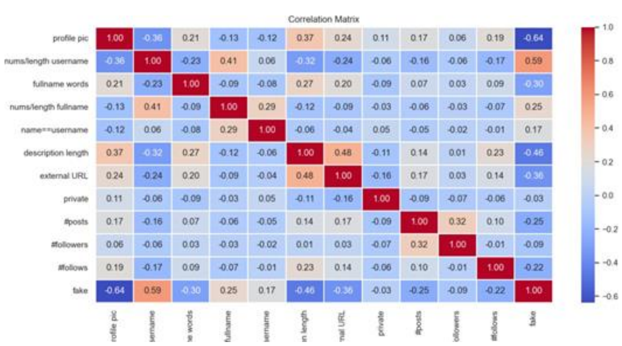
Recall (0.867925) is lower compared to Gradient Boosting and Random Forest, suggesting it might miss more positive cases.

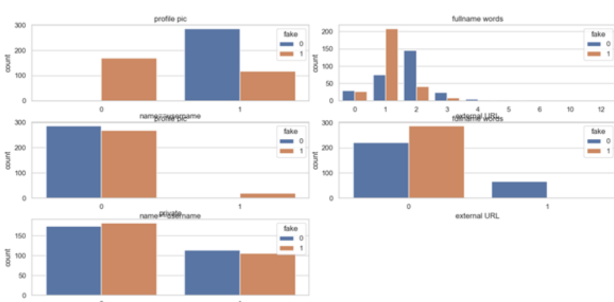
F1 score (0.893204) is the lowest among the three, reflecting that its precision and recall are less balanced compared to the other models.

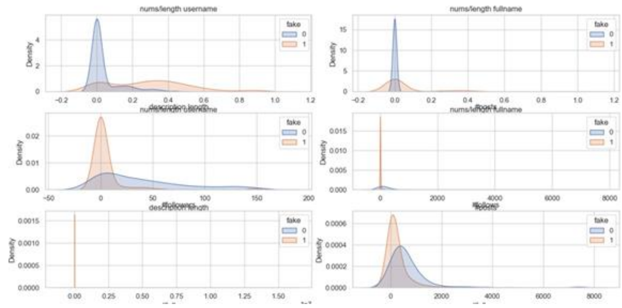
In summary, these metrics help to understand how the models perform in terms of correctly identifying fake profiles (accuracy), avoiding false positives (precision), capturing true positives (recall), and balancing precision and recall (F1 score). These metrics provide a comprehensive view of each model's strengths and weaknesses, allowing stakeholders to choose the most appropriate approach for their specific needs

Various Screenshots









Features and Their Characteristics:

The data columns include a mix of key profile characteristics used to distinguish between genuine and fake accounts:

profile pic: Indicates whether the profile has a picture. nums/length username: Proportion of numbers in the username. fullname words: Count of words in the full name.

nums/length fullname: Proportion of numbers in the full name. name==username: Whether the full name matches the username. description length: The length of the profile description.

external URL: Presence of external URLs on the profile. private: Indicates whether the profile is private.

Social metrics: Number of posts (#posts), followers (#followers), and follows (#follows).

fake: The target variable indicating whether the profile is fake (1 for fake, 0 for genuine).

Overall, Gradient Boosting achieved the highest performance across all metrics (accuracy,
precision, recall, and F1 Score), indicating its effectiveness in distinguishing between fake and genuine profiles. Random Forest and Support Vector Machine also demonstrated high performance, suggesting that they can be suitable alternatives in various contexts.

VI. CONCLUSION

Conclusion

The study's primary focus was on developing and assessing machine learning algorithms to detect fake social media profiles. By examining features like profile pictures, usernames, full names, description lengths, external URLs, privacy settings, and metrics such as the number of posts, followers, and follows, the research aimed to create an effective model for distinguishing between genuine and fake profiles. After comparing three popular machine learning algorithms—Gradient Boosting, Random Forest, and Support Vector Machine—the study found that Gradient Boosting consistently delivered the best performance across various metrics, including accuracy, precision, recall, and F1 score. This algorithm's sequential learning approach, which corrects the mistakes of previous iterations, contributed to its superior performance. The conclusions drawn from this study suggest that Gradient Boosting is a robust and reliable algorithm for detecting fake social media profiles. Its ability to handle complex relationships and correct errors during the learning process gives it an edge over other algorithms. As a result, it can be employed as a key component in combating misinformation and ensuring the authenticity of social media content.

REFERENCE

1. S. Wang Shu and H. Liu, "Understanding User Profiles on social media for Fake News Detection", 2018 IEEE Conference on Multimedia Information Processing and Retrieval (MIPR), pp. 430-435, 2018.
2. Shalinda Adikari and Kaushik Dutta, "Identifying fake profiles in LinkedIn", arXiv preprint, 2020.
3. S. Gurajala, J.S. White, B. Hudson and J.N. Matthews, "Fake Twitter accounts: profile characteristics obtained using an activity-based pattern detection approach", Julyin Proceedings of the 2015 International Conference on social media & Society, pp. 1-7, 2015.
4. Jyoti Kaubiyal and Ankit Kumar Jain, "A feature-based approach to detect fake profiles in Twitter", Proceedings of the 3rd International Conference on Big Data and Internet of Things, pp. 135-139, 2019.
5. F.A. Ozbay and B. Alatas, "Fake news detection within online social media using supervised artificial intelligence algorithms", Physical A: Statistical Mechanics and its Applications, vol. 540, pp. 123174, 2020.
6. A V. N. Reddy and C. Phanikrishna, "Contour tracking based knowledge extraction and object recognition using deep learning neural networks", Proceedings on 2nd International Conference on Next Generation Computing Technologies in 2016 NGCT 2016, pp. 352-354, 2016.
7. N. and M. Smruthi, "Hybrid Scheme for Detecting Fake Accounts in Facebook", (IJRTE)International Journal of Recent Technology and Engineering, vol. 7, no. 5S3, 2019, ISSN 2277-3878.
8. V. Kavitha, G. H. Kumar, S. Kumar, and M. Harish, "Churn prediction of customer in telecom industry using machine learning algorithms," International Journal of Engineering Research and Technology, vol. 9, no. 5, pp. 181–184, 2020.
9. G. Jiao and H. Xu, "Analysis and comparison of forecasting algorithmsfor telecom customer churn," Journal of Physics: Conference Series, vol. 1881, no. 3, Article ID 032061, 2021. Q. Zhu, X. Yu, Y. Zhao, and D. Li, "Customer churn prediction based on LASSO and random forest models," IOP conference series materials science and engineering, vol. 631, no. 5, Article ID 052008, 2019.
10. K. G. Li and B. P. Marikannan, "Hyperparameters tuning and model comparison for telecommunication customer churn predictive models," in Proceedings of the Paper presented at the 3rd Global Conference on Computing and Media Technology, Kuala Lumpur, Malaysia, July 2020.
11. T. Wang, X. Wang, R. Ma, X. Li, and J. Ruan, "Random forest-bayesian optimization for product quality prediction with large-scale dimensions in process industrial cyber–physical systems," IEEE Internet of Things Journal, vol. 7, no. 9, pp. 8641–8653, 2020.
12. L. Yang and A. Shami, "On hyperparameter optimization of machine learning algorithms:

theory and practice," Neurocomputing, vol. 415, pp. 295–316, 2020.

13. N. Hoque, D. K. Bhattacharyya, and J. K. Kalita, "A mutual information-based feature selection method," Expert Systems with Applications, vol. 41, no. 14, pp. 6371–6385, 2014.
14. I. AlShourbaji, N. Helian, Y. Sun, and M. Alhameed, "Anovel HEOMGA approach for class imbalance problem in the application of customer churn prediction," SN Computer Science, vol. 2, no. 6, pp. 1–12, 2021.
15. H. Jain, A. Khunteta, and S. Srivastava, "Churn prediction in telecommunication using logistic regression and logit boost," Procedia Computer Science, vol. 167, pp. 101–112, 2020.
16. S. Wang, L. L. Minku, and X. Yao, "Online class imbalance learning and its applications in fault detection," International Journal of Computational Intelligence and Applications, vol. 12, no. 4, Article ID 1340001, 2013.
17. F. Zhao and Z. Yao, "Predicting the voluntary donation to online content creators," Industrial Management & Data Systems, vol. 120, no. 10, pp. 1941–1957, 2020.
18. A. Onan, S. Korukoğlu, and H. Bulut, "Ensemble of keyword extraction methods and classifiers in text classification," Expert Systems with Applications, vol. 57, pp. 232–247, 2016.
19. Onan, "An ensemble scheme based on language function analysis and feature engineering for text genre classification," Journal of Information Science, vol. 44, no. 1, pp. 28–47, 201