Fake Profile Identification using Machine Learning

Dissertation submitted to

Shri Ramdeobaba College of Engineering & Management, Nagpur

in partial fulfillment of requirement for the award of degree of

Bachelor of Technology (B.Tech)

In

COMPUTER SCIENCE AND ENGINEERING

By

Archit Mahule (25)

Arun Bhangdiya (26)

Dev Gupta (32)

Karan Agrawal (39)

Guide

Prof. Bhagyashree S. Madan

Department of Computer Science and Engineering

Shri Ramdeobaba College of Engineering & Management, Nagpur 440 013

(An Autonomous Institute affiliated to Rashtrasant Tukdoji Maharaj Nagpur University Nagpur)

April 2024

Fake Profile Identification using Machine Learning

Dissertation submitted to

Shri Ramdeobaba College of Engineering & Management, Nagpur

in partial fulfillment of requirement for the award of degree of

Bachelor of Technology (B.Tech)

In

COMPUTER SCIENCE AND ENGINEERING

By

Archit Mahule (25)

Arun Bhangdiya (26)

Dev Gupta (32)

Karan Agrawal (39)

Guide

Prof. Bhagyashree S. Madan

Department of Computer Science and Engineering

Shri Ramdeobaba College of Engineering & Management, Nagpur 440 013

(An Autonomous Institute affiliated to Rashtrasant Tukdoji Maharaj Nagpur University Nagpur)

April 2024

SHRI RAMDEOBABA COLLEGE OF ENGINEERING MANAGEMENT, NAGPUR

(An Autonomous Institute affiliated to Rashtrasant Tukdoji Maharaj Nagpur University Nagpur)

Department of Computer Science and Engineering

CERTIFICATE

This is to certify that the Thesis on **"Fake Profile Identification using Machine Learning"** is a Bonafide work of Archit Mahule, Arun Bhangdiya, Dev Gupta and Karan Agrawal, submitted to the Rashtrasant Tukdoji Maharaj Nagpur University, Nagpur in partial fulfillment of the award of a Degree of Bachelor of Technology (B.Tech), in Computer Science and Engineering. It has been carried out at the Department of Computer Science and Engineering, Shri Ramdeobaba College of Engineering and Management, Nagpur during the academic year 2023-2024.

Date:

Place: Nagpur

Prof. Bhagyashree S. Madan Project Guide Department of Computer Science and Engineering	Dr. R. Hablani H.O.D Department of Computer Science and Engineering

Dr. R. S. Pande
Principal

DECLARATION

We hereby declare that the thesis titled "**Fake Profile Identification using Machine Learning**" submitted herein, has been carried out in the Department of Computer Science and Engineering of Shri Ramdeobaba College of Engineering and Management, Nagpur. The work is original and has not been submitted earlier as a whole or part for the award of any degree/diploma at this or any other institution / University.

Date:

Place: Nagpur

Name	Roll No.	Signature
Archit Mahule	**25**	
Arun Bhangdiya	**26**	
Dev Gupta	**32**	
Karan Agrawal	**39**	

APPROVAL SHEET

This report entitled " **Fake Profile identification using Machine Learning** " by Archit Mahule (25), Arun Bhangdiya (26), Dev Gupta (32) and Karan Agrawal (39) is approved for the degree of Bachelor of Technology (B.Tech).

Prof. Bhagyashree S. Madan Project Guide	External Examiner

Dr. R. Hablani
H.O.D, CSE

Date:

Place: Nagpur

ACKNOWLEDGEMENTS

The project is a combined effort of a group of individuals who synergize to contribute towards the desired objectives. Apart from our efforts by us, the success of the project shares an equal proportion with the engagement and guidance of many others. We take this opportunity to express our gratitude to the people who have been instrumental in the successful completion of this project.

We would like to show our greatest appreciation towards Dr. Khushboo Khurana for providing us with the facilities for completing this project and for their constant guidance.

We would like to express our deepest gratitude to Dr. Ramchand Hablani, Head, Department of Computer Science and Engineering, RCOEM, Nagpur for providing us the opportunity to embark on this project.

Finally, extend our gratitude to all the faculty members of the CSE department who have always been so supportive and provided resources needed in our project development. We are very grateful to all of them who have unconditionally supported us throughout the project.

Date:

Projectees -

Archit Mahule

Arun Bhangdiya

Dev Gupta

Karan Agrawal

ABSTRACT

In the realm of social media, particularly on platforms like Instagram, the proliferation of fake profiles poses a significant challenge. In this project, we address the task of detecting fake profiles on Instagram through the lens of machine learning techniques. Our methodology revolves around employing three distinct classifiers: Gradient Boosting, Random Forest, and Support Vector Machine (SVM).

We begin by loading and exploring the dataset, 'instagram.csv', where we examine features and their distributions, visualize correlations, and identify outliers. To enhance model performance, we introduce novel features such as engagement ratio, social interaction index, profile completeness score, and average description word length. Missing or infinite values are handled meticulously, and outliers are addressed through data preprocessing techniques including log transformation.

Model building and evaluation involve implementing the aforementioned classifiers individually, scrutinizing their accuracy, precision, recall, and F1 score. The ensemble model, formed through a Voting Classifier with the base models, further bolsters our predictive capabilities. Evaluation metrics on both training and test sets affirm the efficacy of our approach.

Moreover, we extend our project into a Flask environment, integrating frontend and backend elements for seamless usability. Leveraging the Instaloader library, we incorporate real-time data scraping functionality, enabling the model to predict on user data fetched from Instagram usernames.

Our results showcase promising outcomes, with the ensemble model achieving a notable test set accuracy of 93.10%. This project not only contributes to the ongoing efforts in

combating online fraudulence but also underscores the potential of machine learning in safeguarding social media integrity. contributes to the ongoing efforts in combating online fraudulence but also underscores the potential of machine learning in safeguarding social media integrity.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLE

CHAPTER 1
INTRODUCTION

1.1 Problem Definition

Social media platforms have revolutionized the way individuals communicate, share information, and engage with others online. Among these platforms, Instagram stands out as one of the most popular platforms for visual content sharing, boasting over a billion active users worldwide. However, along with its widespread adoption comes the pervasive issue of fake profiles. These profiles, often created with malicious intent, undermine the authenticity of user interactions, compromise user privacy, and facilitate various forms of online fraud.

The problem of fake profiles on Instagram manifests in multiple ways, including but not limited to:

- Identity theft: Fake profiles may impersonate real individuals, stealing their identities and potentially causing harm to their reputation or financial standing.
- Spamming: Fake profiles may engage in spamming activities, inundating users with unsolicited messages, links, or advertisements.
- Misinformation dissemination: Fake profiles may spread false or misleading information, contributing to the proliferation of misinformation and undermining trust in online content.

Identifying and mitigating the presence of fake profiles on Instagram is thus a critical challenge that requires effective detection mechanisms and proactive measures to safeguard user trust and platform integrity.

1.2 Motivation

The prevalence of fake profiles on Instagram not only undermines user trust but also poses significant challenges for platform administrators, businesses, and individual users. Detecting and removing fake profiles manually is a daunting and time-consuming task, often requiring considerable resources and expertise. Moreover, the rapid evolution

of fake profile creation techniques necessitates innovative and automated approaches for detection and mitigation.

The motivation behind this project stems from the urgent need to address the problem of fake profiles on Instagram through advanced machine learning techniques. By leveraging the power of data-driven algorithms, we aim to develop a robust and scalable solution that can effectively identify and flag fake profiles, thereby enhancing the overall security and trustworthiness of the platform.

Furthermore, by raising awareness of the risks associated with fake profiles and demonstrating the feasibility of automated detection methods, we hope to empower users and platform administrators to take proactive measures in combating online fraudulence.

1.3 Overview

In this project, we embark on a comprehensive exploration of fake profile detection on Instagram using machine learning methodologies. Our approach encompasses several key stages, including data exploration, feature engineering, model building, evaluation, and deployment. By following a structured methodology, we aim to systematically tackle the complexities of fake profile detection and develop a robust solution that can adapt to evolving threats and challenges.

The project's scope includes:

- Exploring the characteristics and patterns associated with fake profiles on Instagram through thorough data analysis.
- Developing novel features and metrics to enhance the discriminatory power of machine learning models in distinguishing between fake and genuine profiles.
- Building and evaluating multiple machine learning classifiers, including Gradient Boosting, Random Forest, and Support Vector Machine (SVM), to identify the most effective approach for fake profile detection.
- Leveraging ensemble learning techniques to combine the strengths of individual classifiers and improve overall detection performance.

- Integrating real-time data scraping capabilities using the Instaloader library to enable the model to predict on user data fetched from Instagram usernames.
- Implementing a user-friendly Flask environment to facilitate the deployment and usage of the developed solution.

Through these efforts, we aim to contribute to the ongoing discourse on social media security and promote a safer and more trustworthy online environment for users worldwide.

1.4 Objectives

The primary objectives of this project are as follows:

1. Explore the multifaceted problem of fake profile detection on Instagram and understand its implications for user trust and platform integrity.
2. Develop a comprehensive understanding of the characteristics and behaviors exhibited by fake profiles through thorough data analysis and exploration.
3. Design and implement novel features and metrics to enhance the discriminatory power of machine learning models in identifying fake profiles.
4. Build, train, and evaluate multiple machine learning classifiers, including Gradient Boosting, Random Forest, and Support Vector Machine (SVM), to identify the most effective approach for fake profile detection.
5. Investigate ensemble learning techniques to combine the predictions of individual classifiers and improve overall detection performance.
6. Integrate real-time data scraping functionality using the Instaloader library to enable the model to predict on user data fetched from Instagram usernames.
7. Implement a user-friendly Flask environment to facilitate the deployment and usage of the developed solution, ensuring accessibility and usability for a wide range of users.

By achieving these objectives, we aim to develop a robust and scalable solution for fake profile detection on Instagram, contributing to the broader goal of promoting online safety and trust in social media platforms.

CHAPTER 2
LITERATURE REVIEW

In recent years, the detection of fraudulent profiles on online social networks (OSNs) has garnered significant attention from researchers worldwide. This chapter provides a comprehensive overview of the literature related to the identification and prevention of suspicious accounts on social media platforms.

2.1 Logistic Regression with Gradient Descent (LR-GD)

Eswara Sai Raja et al. (2023) proposed the use of Logistic Regression with Gradient Descent (LR-GD) for detecting suspicious accounts in OSNs. Utilizing a dataset comprising 969 profiles, their technique achieved an impressive accuracy of 92.7% and a false positive rate of 0.5%. The LR-GD model demonstrated effectiveness in distinguishing between genuine and fraudulent profiles, highlighting its potential for enhancing online security [1].

2.2 XGBoost for Fraudulent Profile Detection

K. Harish et al. (2023) investigated the application of XGBoost, an efficient gradient boosting algorithm, for detecting fraudulent profiles on social media platforms. Their study reported an accuracy rate of 94.2% for the XGBoost model, demonstrating its efficacy in identifying and mitigating cyber security risks posed by fraudulent accounts. Additionally, the authors suggested future research directions, including extending the technique to other social media platforms and developing real-time detection models [2].

2.3 Natural Language Processing (NLP) and Machine Learning

Dr. K. Smita et al. (2023) proposed a system that combines Natural Language Processing (NLP) and machine learning techniques for identifying Twitter spammers and false users. Their approach achieved a commendable accuracy rate of 92.4%, underscoring the effectiveness of NLP-based methods in combating fraudulent activities

on social media platforms. Furthermore, the authors outlined potential areas for further research, such as detecting fake news and analyzing social network dynamics [3].

2.4 Artificial Neural Networks (ANNs) for False Profile Detection

Partha Chakraborty et al. (2022) explored the use of Artificial Neural Networks (ANNs) for detecting and preventing false profiles on social media platforms. Their study demonstrated the efficacy of a 6-layered ANN model, achieving an accuracy rate of 91.8%. The authors emphasized the importance of real-time applications, the integration of additional parameters, and the combination of different machine learning models as avenues for future investigation [4].

2.5 Machine Learning Techniques for Cloned Profile Detection

Ajith M. and M. Nirmala (2022) proposed a technique utilizing machine learning algorithms for the detection of cloned and bogus social media profiles, with a focus on Twitter. Their research highlighted the successful application of machine learning techniques in achieving this objective. Additionally, the authors suggested leveraging natural language processing (NLP) approaches to further enhance detection accuracy [5].

2.6 Machine Learning Algorithms for Fake Profile Identification

M. Mamatha et al. (2021) investigated the use of machine learning algorithms for identifying fake profiles on social media platforms. While their method demonstrated success in detecting phony profiles, the authors acknowledged the challenge of achieving extremely high accuracy. They proposed avenues for future research, including refining the model, exploring new features, and implementing real-time detection mechanisms for improved accuracy [6].

In summary, the literature survey highlights a diverse range of approaches and techniques employed by researchers to address the challenge of fraudulent profile detection on online social networks. These studies underscore the significance of machine learning, natural language processing, and artificial intelligence in combating cyber security threats and safeguarding the integrity of social media platforms.

CHAPTER 3

METHODOLOGY

3.1 Dataset

The dataset utilized in this project comprises data from over 500 Instagram profiles, encompassing 12 distinct features. These features offer a comprehensive view of user behavior, profile characteristics, and engagement metrics, which are pivotal for the detection of fake profiles. Noteworthy attributes include numeric metrics such as '#posts', '#followers', and '#follows', providing quantitative measures of user activity and popularity. The target variable, denoted as 'fake', categorizes profiles as either genuine (0) or fake (1), forming the foundation for supervised learning tasks.

Table 1: Dataset Features Description.

Sr. No.	Attribute	Description	Type	Why it's Useful
1	profile pic	Indicates whether the account has a profile picture or not.	presence(1)/absence(0)	Fake accounts may lack a profile picture to avoid detection.
2	nums/length username	Proportion of numerical characters in the username to its total length.	ratio	High ratio might indicate automatically generated usernames for fake accounts.
3	fullname words	The number of words in the user's full name.	number	Suspicious accounts might use generic one-word names or incomplete names.
4	nums/length fullname	Similar to username ratio, but for the full name.	ratio	High values could suggest fake or stolen names with random characters.
5	name==username	Whether the username and full name are identical.	boolean	True could indicate low effort account creation for bots or fake profiles.
6	description length	Character count of the user's bio.	number	Extremely short or long descriptions might be suspicious for automated generation.
7	external URL	Presence of an external link in the bio.	boolean	Suspicious accounts might use it for spamming or phishing.

8	private	Whether the account is private or public.	boolean	Private accounts with limited activity could be suspicious.
9	#posts	Total number of posts made by the user.	number	Very low or high post counts compared to followers might be suspicious.
10	#followers	Number of users following the account.	number	Large discrepancies between followers and activity could be suspicious (bought followers).
11	#follows	Number of users the account follows.	number	Following many accounts with low engagement could be suspicious (bot activity).
12	fake	The target variable indicates whether the account is classified as genuine or fake.	label	This is the variable your model will predict based on the other attributes.

The target variable distribution of real and fake profiles are as given in fig 1.

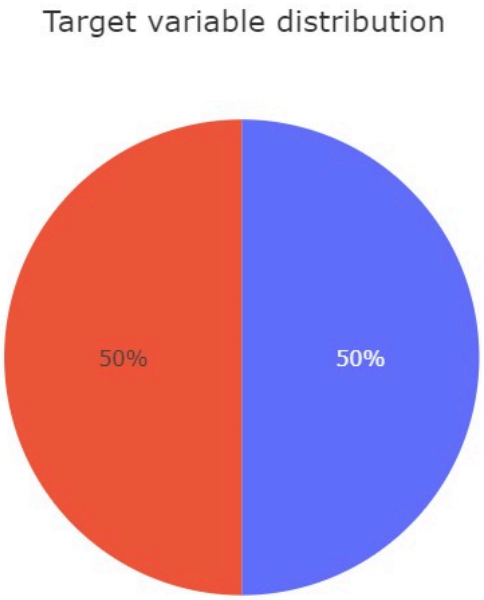


Fig. 1. Target Variable Distribution of Dataset

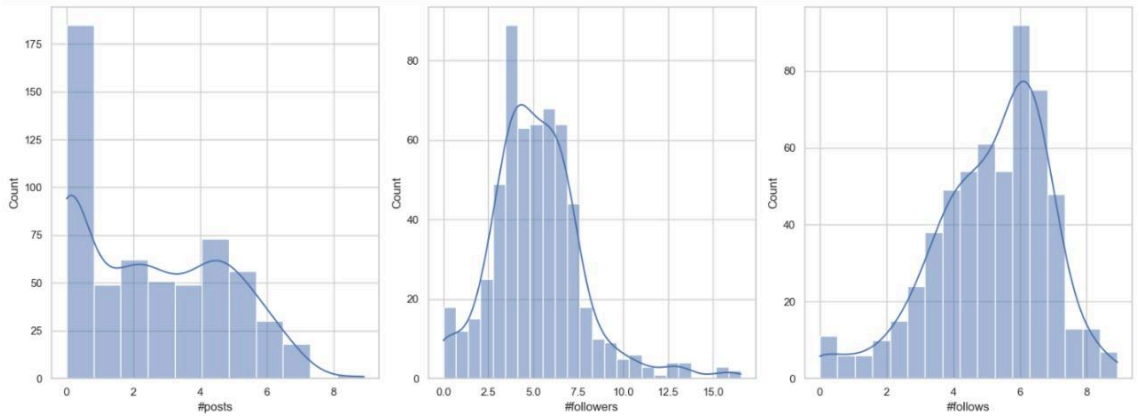


Fig. 2. Numerics Metrics like Post, Followers and Follows

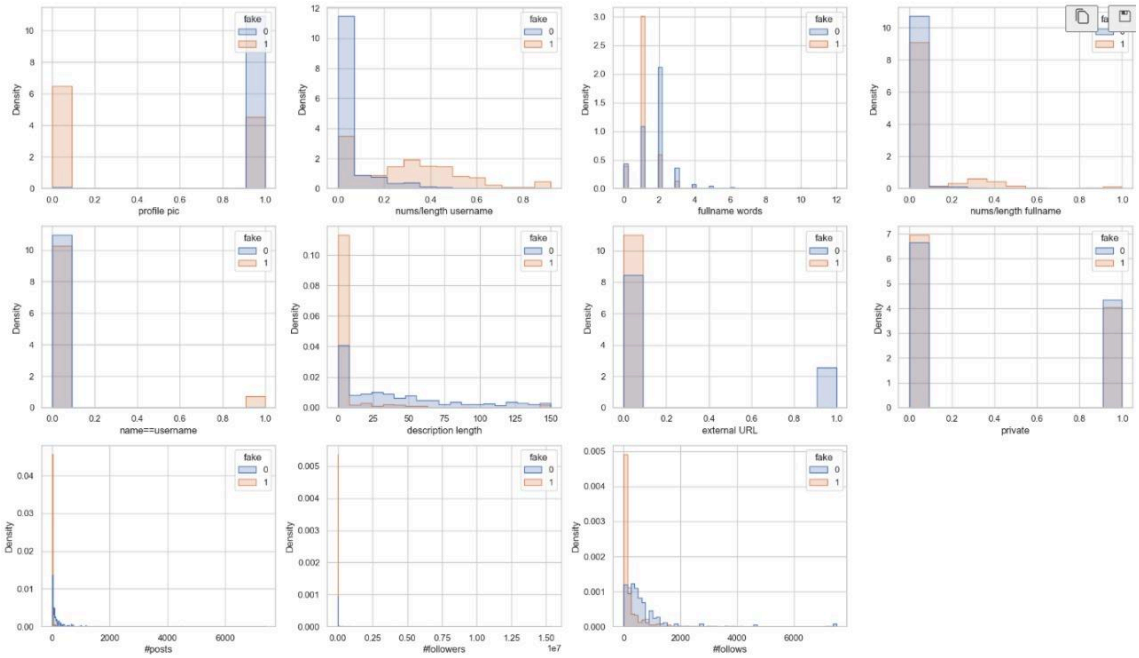


Fig. 3. Visualization of data about fake and real Instagram profiles

3.2 Proposed Methodology

The proposed methodology adopts a holistic approach to fake profile detection, encompassing data preprocessing, feature engineering, model building, evaluation, and deployment. Key components of the methodology include:

1. Data Preprocessing: This stage involves handling missing values, identifying outliers, and transforming features to ensure data integrity and suitability for modeling.

2. Feature Engineering: In addition to existing features, novel features such as engagement ratio, social interaction index, and profile completeness score are introduced. These features offer deeper insights into user behavior and profile characteristics, enhancing the discriminatory power of the models.
    a. Average Description Length: This feature assesses the level of engagement per post, providing valuable insights into the quality of user interactions.

$$Average\ Description\ Length\ = \frac{Total\ length\ of\ description}{total\ number\ of\ words\ in\ description}$$

    b. Engagement Ratio: Calculated as the ratio of followers_count to statuses_count, this metric offers a measure of the user's engagement level relative to their follower count.

$$Engagement\ Ratio\ = \frac{Followers\ Count}{Statuses\ Count}$$

    c. Social Interaction Index: Capturing the overall level of social interaction and activity, this index offers a comprehensive view of profile engagement.

$$Social\ Interaction\ Index = Followers\ Count + Friends\ Count + Favourites\ Count$$

    d. Profile Completeness Score: Evaluating the completeness of profile information, including the presence of a URL, profile image, and bio, this score aids in assessing profile authenticity.

$$Profile\ Completeness\ Score = Profile\ Picture + External\ URL + (Description\ Length > 0)$$

3. Model Building: Three classification models are employed: Gradient Boosting, Random Forest, and Support Vector Machine (SVM). Each model offers unique advantages in tackling the fake profile detection problem:
    a. Gradient Boosting: This ensemble learning technique builds sequential weak learners, each correcting the errors of its predecessor. It is known for its robustness and ability to handle complex datasets.

b. Random Forest: Based on the concept of decision trees, Random Forest constructs multiple trees and aggregates their predictions to produce a final output. It is well-suited for handling large datasets and provides insights into feature importance.
c. Support Vector Machine (SVM): SVM operates by finding the hyperplane that best separates classes in a high-dimensional space. It is effective in handling both linearly and nonlinearly separable data, making it a versatile choice for classification tasks.

4. Model Evaluation: The performance of each model is evaluated using a range of metrics including accuracy, precision, recall, and F1 score. These metrics provide insights into the model's ability to correctly classify genuine and fake profiles, as well as its overall predictive power.
5. Ensemble Learning: Ensemble learning techniques, such as Voting Classifier, are employed to combine the predictions of multiple base models and improve overall detection performance. By leveraging the strengths of individual classifiers, ensemble learning enhances the robustness and reliability of the detection system.

3.3 Architecture

3.3.1 Classification Models Used

The classification models utilized in this project, namely Gradient Boosting, Random Forest, and Support Vector Machine (SVM), offer distinct approaches to fake profile detection:

1. Gradient Boosting: This model iteratively builds an ensemble of weak learners, each attempting to correct the errors made by its predecessor. By combining multiple weak learners, Gradient Boosting creates a strong predictive model capable of handling complex relationships within the data.
2. Random Forest: Constructed from multiple decision trees, Random Forest generates predictions by aggregating the outputs of individual

trees. This model is renowned for its scalability, robustness, and ability to handle high-dimensional datasets.

3. Support Vector Machine (SVM): SVM operates by identifying the hyperplane that best separates classes in a high-dimensional feature space. It is particularly effective in scenarios where the data is not linearly separable, thanks to its ability to map data into higher-dimensional spaces.

CHAPTER 4
IMPLEMENTATION

In this chapter, we delve into the detailed implementation process of our system designed for detecting fraudulent profiles on online social networks (OSNs). This involves four primary stages: data collection, data preprocessing, model training, and the development of a web application that allows for real-time profile verification.

4.1 Data Collection

Our first step involved collecting data for training and evaluating our model. We utilized a Python library named Instaloader to extract profile information from Instagram, one of the largest online social networks. We extracted a collection of features for each profile that could possibly be indicative of fraudulent activity. These features include:

- Presence of a profile picture: Fraudulent profiles often lack a personal photo.
- Number of numerical characters and length of the username: A high number of numerical characters or an unusually long username may suggest suspicious activity.
- Number of words, numerical characters, and length of the full name: Similar to the username, these could be potential indicators of a fraudulent profile.
- Whether the username matches the full name: A mismatch may suggest that the profile is fraudulent.
- Length of the description: Short or non-existent descriptions may be a red flag.
- Presence of an external URL: Fraudulent profiles may link to suspicious external sites.
- Account privacy status: Public or private status could be a potential indicator.
- Number of posts, followers, and accounts followed by the user: An unusually high or low number in any of these categories may suggest fraudulent activity.

To enhance the predictive power of our model, we also engineered four new features:

- Engagement ratio: This is the ratio of followers to following, and it can help identify profiles with unnatural follower dynamics.

- Social interaction index: This measure of interaction with other users can reveal patterns typical of fraudulent profiles.
- Profile completeness score: Incomplete profiles may be more likely to be fraudulent.
- Average description word length: Unusually short or long words may be a sign of a bot-generated description.

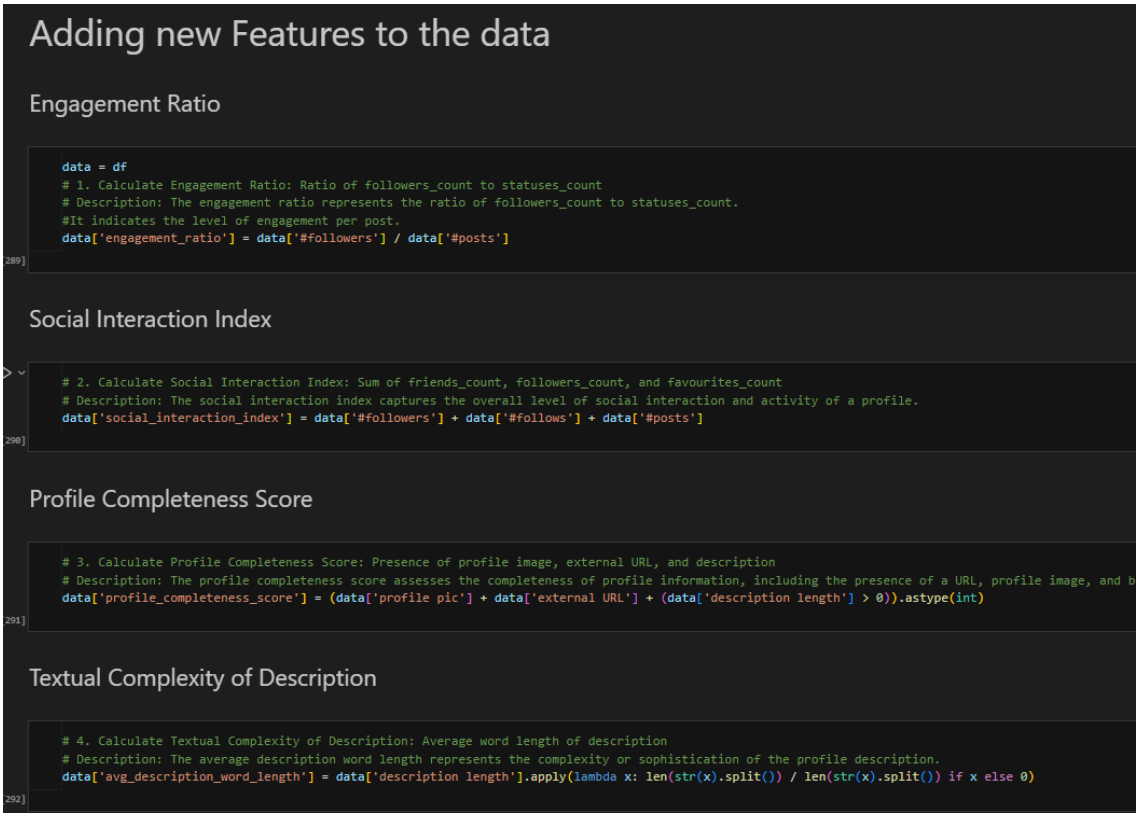
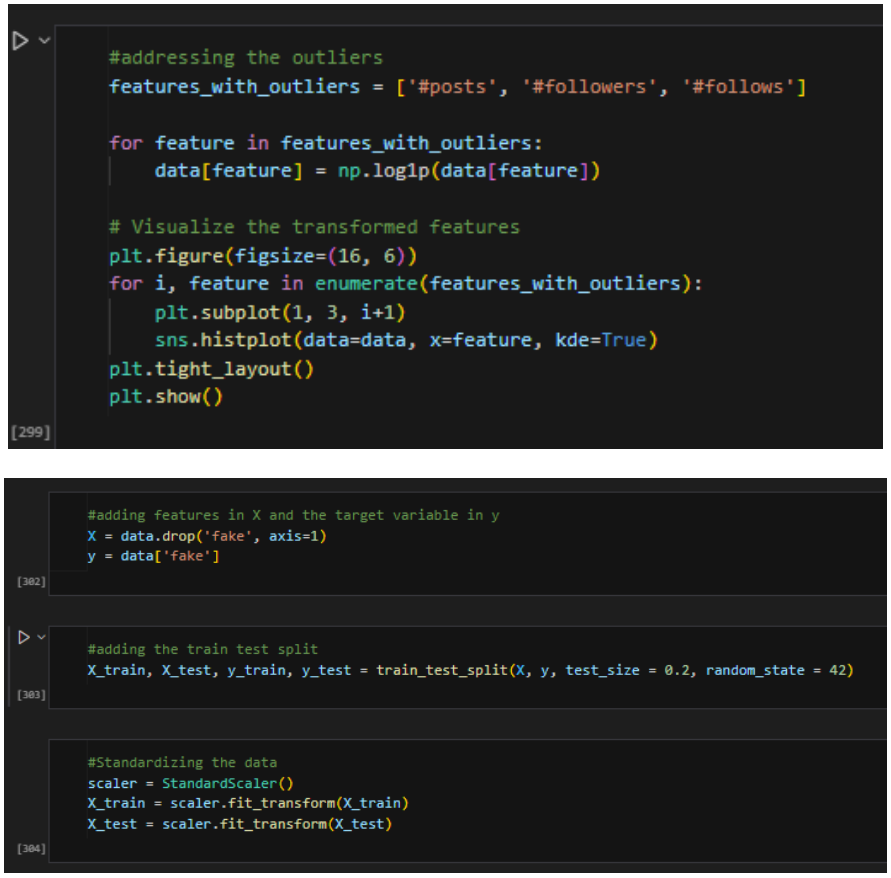


Fig. 4. Adding new features

4.2 Data Preprocessing

Following data collection, we carried out preprocessing on the data to ensure its readiness for the machine learning models. This involved dealing with missing values, outliers, and feature scaling. Missing values, particularly in the engagement ratio feature, were replaced with the mean value of the column to ensure data completeness. Outliers in numerical features such as the number of posts, followers, and accounts followed were addressed using log transformation. This approach handles skewed distribution and improves model performance by reducing the influence of data points that are significantly distant from the rest. We also applied standard scaling to normalize

the data and bring all features to a similar scale, reducing the chance of certain features dominating others due to their magnitude.



```
#addressing the outliers
features_with_outliers = ['#posts', '#followers', '#follows']

for feature in features_with_outliers:
    data[feature] = np.log1p(data[feature])

# Visualize the transformed features
plt.figure(figsize=(16, 6))
for i, feature in enumerate(features_with_outliers):
    plt.subplot(1, 3, i+1)
    sns.histplot(data=data, x=feature, kde=True)
plt.tight_layout()
plt.show()
[299]
```

```
#adding features in X and the target variable in y
X = data.drop('fake', axis=1)
y = data['fake']
[302]

#adding the train test split
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size = 0.2, random_state = 42)
[303]

#Standardizing the data
scaler = StandardScaler()
X_train = scaler.fit_transform(X_train)
X_test = scaler.fit_transform(X_test)
[304]
```

Fig. 5. Data Preprocessing.

4.3 Model Training

With the preprocessed data, we proceeded to train multiple machine learning models to identify fraudulent profiles. The models we employed include:

- Gradient Boosting Classifier: This model builds new trees which complement the already built ones, reducing the errors in the predictions.
- Random Forest Classifier: This model creates an ensemble of decision trees and averages their results, reducing the effect of individual errors.
- Support Vector Machine (SVM) Classifier: This model finds the hyperplane in an N-dimensional space that distinctly classifies the data points.

We also employed a voting classifier ensemble technique. This technique combines the predictions from multiple base models to improve overall performance. It essentially

'votes' on the most likely prediction, drawing on the strengths of each individual model. Both hard voting (majority voting) and soft voting (weighted probabilities) strategies were explored to determine the optimal ensemble approach.

```python
# For each model, running a for loop
for model_name, model in models.items():
    # Train the model
    model.fit(X_train, y_train)

    # Make predictions using the test set
    y_pred = model.predict(X_test)

    # Calculating the performance metrics:
    accuracy = accuracy_score(y_test, y_pred)
    precision = precision_score(y_test, y_pred)
    recall = recall_score(y_test, y_pred)
    f1 = f1_score(y_test, y_pred)

    # Results of the model
    results["Model"].append(model_name)
    results["Accuracy"].append(accuracy)
    results["Precision"].append(precision)
    results["Recall"].append(recall)
    results["F1 Score"].append(f1)

    # Print results for individual models
    print(f"{model_name}:")
    print(f"Test set accuracy: {accuracy}\n")

# Converting the results into a DataFrame for easier viewing
results_df = pd.DataFrame(results)
print("Results for Individual Models:")
print(results_df)


# Ensemble Voting Classifier
model_list = [(model_name, model) for model_name, model in models.items()]

# Initialize a VotingClassifier with the models
voting_classifier = VotingClassifier(estimators=model_list, voting='hard')

# Train the ensemble model
voting_classifier.fit(X_train, y_train)

# Make predictions using the training set
y_train_pred_ensemble = voting_classifier.predict(X_train)

# Calculate the training accuracy for the ensemble model
train_accuracy_ensemble = accuracy_score(y_train, y_train_pred_ensemble)

# Make predictions using the test set
y_test_pred_ensemble = voting_classifier.predict(X_test)

# Calculate the test accuracy for the ensemble model
test_accuracy_ensemble = accuracy_score(y_test, y_test_pred_ensemble)

# Print the results for the ensemble model
print("Ensemble Model:")
print(f"Training set accuracy: {train_accuracy_ensemble}")
print(f"Test set accuracy: {test_accuracy_ensemble}")
```

Fig. 6. Model Training.

4.4 Web Application Development

Lastly, a web application was developed using the Flask framework to provide users with a platform for profile verification in real-time. Users can input an Instagram username, and the application retrieves the corresponding profile data using the Instaloader library. The trained ensemble model then processes this data, predicting whether the profile is genuine or fraudulent based on the extracted features. The result is

displayed to the user, clearly indicating whether the profile has been classified as real or fake, aiding in the detection and avoidance of fraudulent profiles.

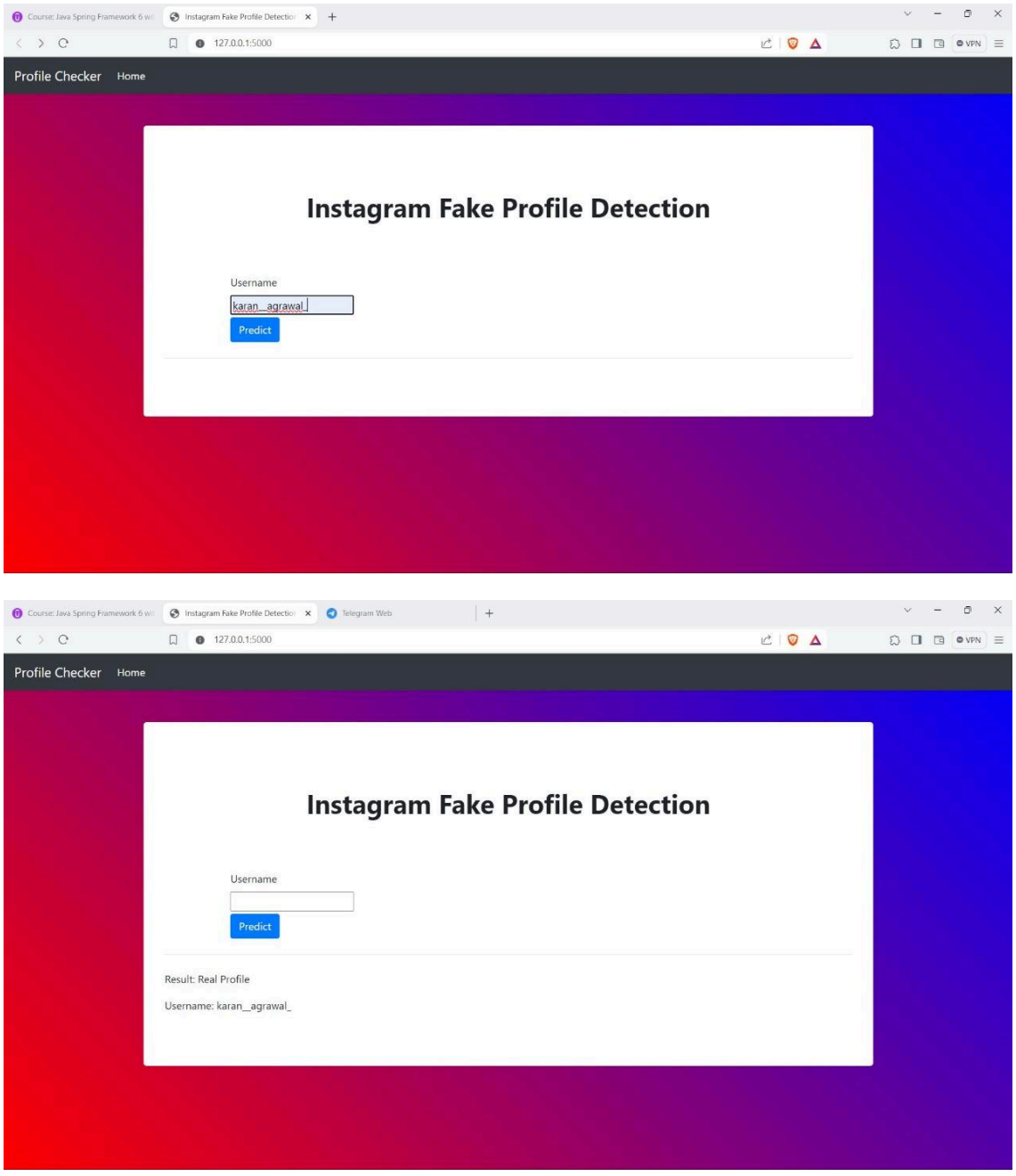


Fig. 7. Testing a Real Profile

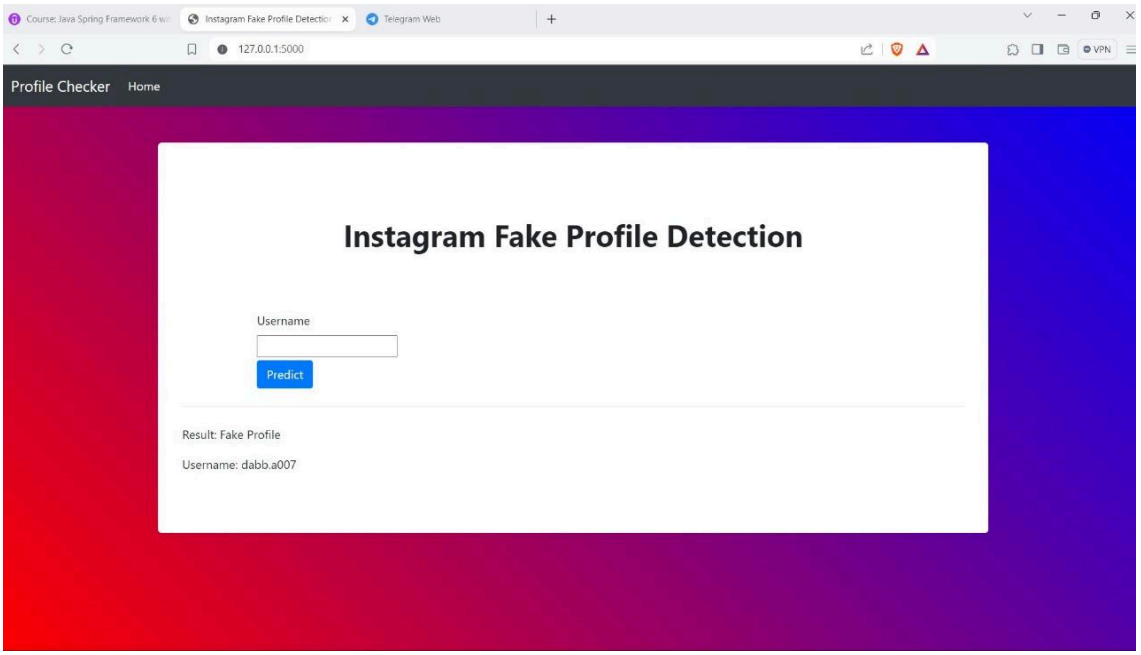


Fig. 8. Testing a Fake Profile

4.5 System Specification

The system architecture is designed to accommodate the development and deployment of machine learning models within a user-friendly web environment. Leveraging Python for backend development, Flask for web application development, and Bootstrap for frontend design, the system ensures seamless integration of machine learning algorithms into a user-friendly interface.

4.6 Technology Stack

1. Python:
    - Role in the Project: Python serves as the primary programming language for the entire project, offering a rich ecosystem of libraries and tools for data processing, machine learning, and web development.
    - Usage in the Project: Python is utilized for various tasks including data preprocessing, feature engineering, model building, evaluation, and deployment. Libraries such as pandas, NumPy, and scikit-learn are leveraged for data manipulation and machine learning tasks.

2. Flask:
    - Role in the Project: Flask is a lightweight and flexible web framework for Python, ideal for developing web applications and APIs.
    - Usage in the Project: Flask is utilized to create a user-friendly web interface for the deployment of machine learning models. It facilitates the integration of backend machine learning functionality with frontend design, enabling real-time prediction on Instagram profiles. Flask's routing mechanisms, request handling, and templating engine are leveraged to build the web application.
3. scikit-learn:
    - Role in the Project: scikit-learn is a popular machine learning library in Python, offering a wide range of algorithms and tools for classification, regression, clustering, and more.
    - Usage in the Project: scikit-learn is extensively used for model building, evaluation, and feature engineering tasks. It provides implementations of classification algorithms such as Gradient Boosting, Random Forest, and Support Vector Machine (SVM), which are crucial for fake profile detection. Additionally, scikit-learn's utilities for data preprocessing, cross-validation, and model evaluation are instrumental in building robust machine learning pipelines.
4. Bootstrap:
    - Role in the Project: Bootstrap is a front-end framework for building responsive and visually appealing web interfaces.
    - Usage in the Project: Bootstrap is employed to design and style the user interface of the web application. Its grid system, CSS components, and JavaScript plugins are utilized to create a responsive and intuitive interface for users to interact with the fake profile detection system. Bootstrap's pre-designed components and responsive layout ensure a consistent user experience across different devices and screen sizes.

By leveraging Python, Flask, scikit-learn, and Bootstrap, the project integrates backend machine learning functionality with a user-friendly web interface, enabling users to interact with the fake profile detection system in a seamless and intuitive manner. Each

component of the technology stack plays a crucial role in different facets of the project, contributing to its overall success in combating fake profiles on Instagram.

CHAPTER 5
RESULT AND EVALUATION

5.1 Performance Evaluation of Individual Models

The study evaluates the performance of several individual models, including Support Vector Machine, Gradient Boosting, and Random Forest, alongside an ensemble model. Among these, the Support Vector Machine exhibits the highest test set accuracy, achieving 92.24%. Similarly, the ensemble model performs well, achieving a test set accuracy of 93.10%. The evaluation metrics including test set accuracy, precision, recall, and F1 score for each model are presented in Table 2 below.

Table 2: Evaluation Metrics

Model	Test Set Accuracy	Precision	Recall	F1 Score
Support Vector Machine	0.922414	0.894	0.885792	0.912621
Gradient Boosting	0.913793	0.921569	0.886792	0.903846
Random Forest	0.913793	0.90566	0.98566	0.90566

5.2 Visualization of Model Performance

Figure 1 illustrates the precision, recall, and F1 score of individual models. The graph provides a visual comparison of the performance metrics, showcasing the strengths and weaknesses of each model.

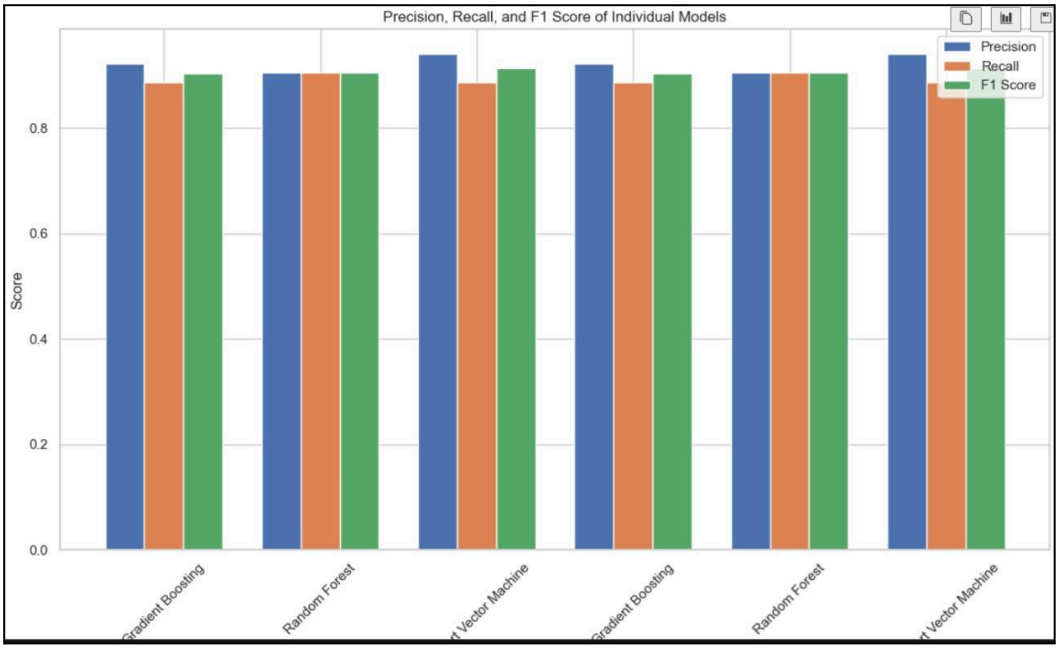


Fig. 9. Precision, Recall, and F1 Score of Individual Models

Additionally, Figure 2 presents a comparative analysis of the accuracy of Support Vector Machine, Gradient Boosting, and Random Forest models. The graph highlights the varying degrees of accuracy achieved by each model, with the Support Vector Machine exhibiting the highest accuracy at 92.24%.

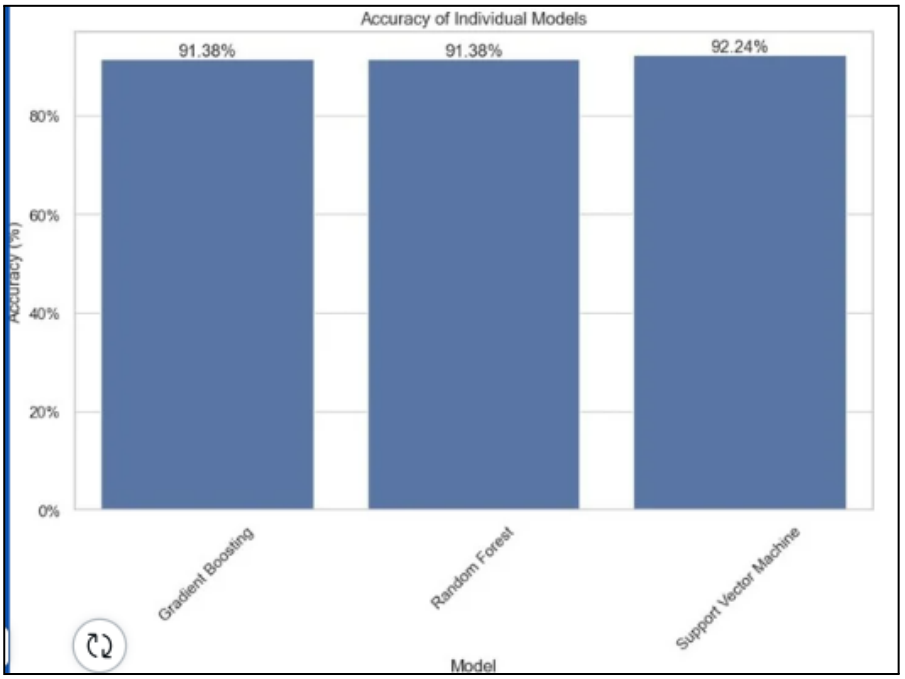


Fig. 10. Accuracy of Individual Models

5.3 Correlation Matrix Analysis

The correlation matrix offers insights into the relationships between various variables in the social media dataset. Positive correlations indicate associations between variables, while negative correlations suggest inverse relationships. Notable findings include:

Positive correlations between the length of a user's profile image and the number of followers, indicating that profiles with longer images tend to have more followers.

Negative correlations between the length of profile images and the likelihood of an account being fake, implying that shorter profile images are associated with fake accounts.

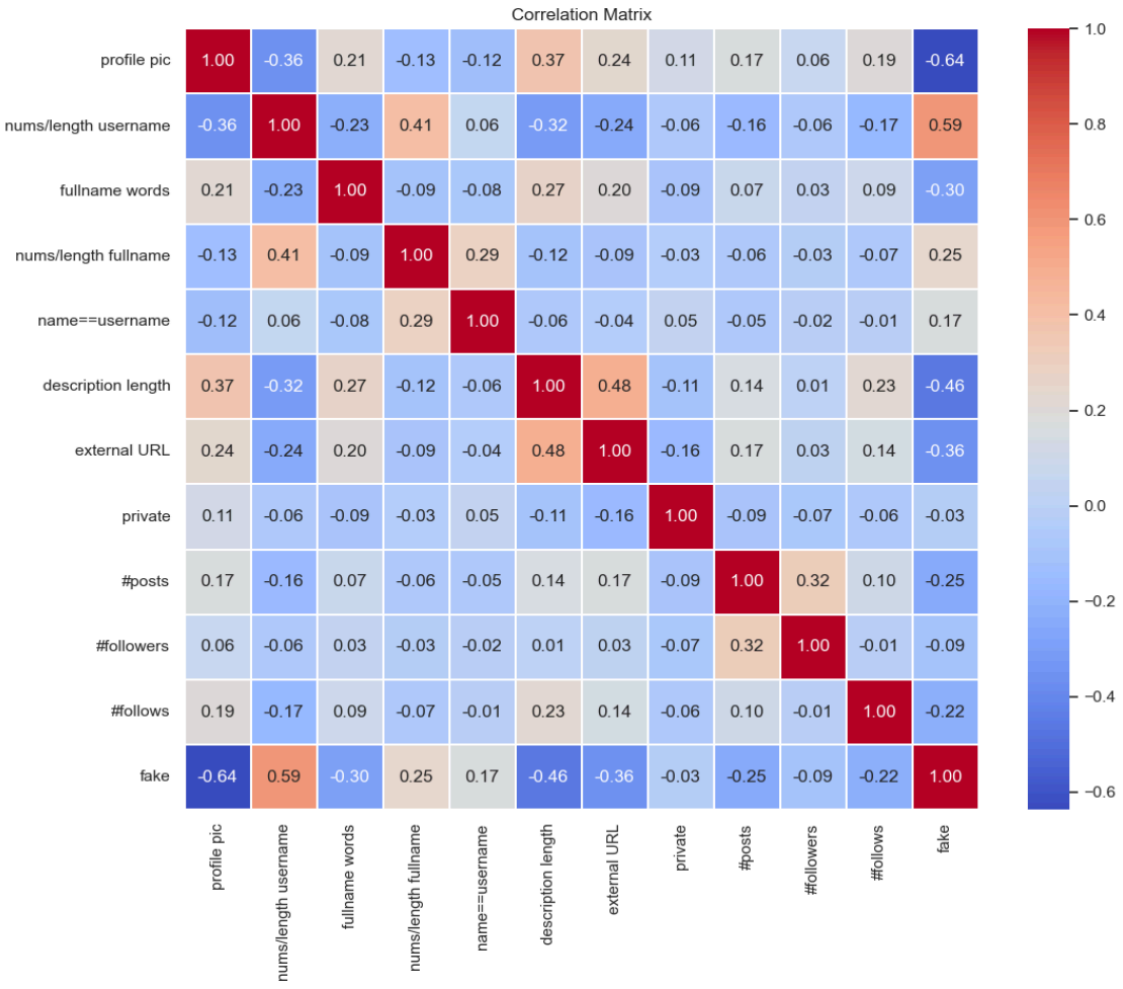


Fig. 11. Correlation Matrix

5.4 ROC Curve Analysis

The ROC curve depicts the performance of the ensemble model (soft voting) on a classification test. The curve illustrates the relationship between the True Positive Rate (TPR) and False Positive Rate (FPR) across different classification thresholds. The ensemble model demonstrates superior performance, with an Area Under the Curve (AUC) of 0.98, indicating exceptional performance in distinguishing between positive and negative instances.

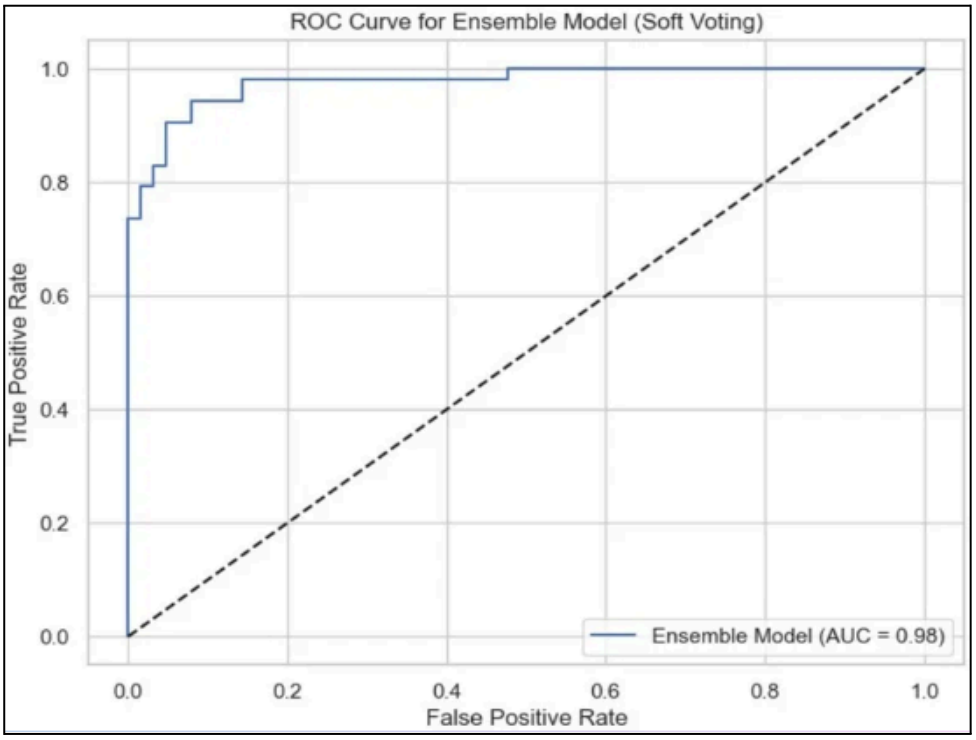


Fig. 12. ROC Curve for Ensemble Model

5.5 Confusion Matrix Analysis

The confusion matrix provides insights into the performance of the ensemble model on binary classification tasks. It depicts the number of correct and incorrect predictions for each class, enabling the calculation of metrics such as accuracy and precision. In this study, the ensemble model demonstrates accurate predictions for both classes, with a notable accuracy of 79.3%.

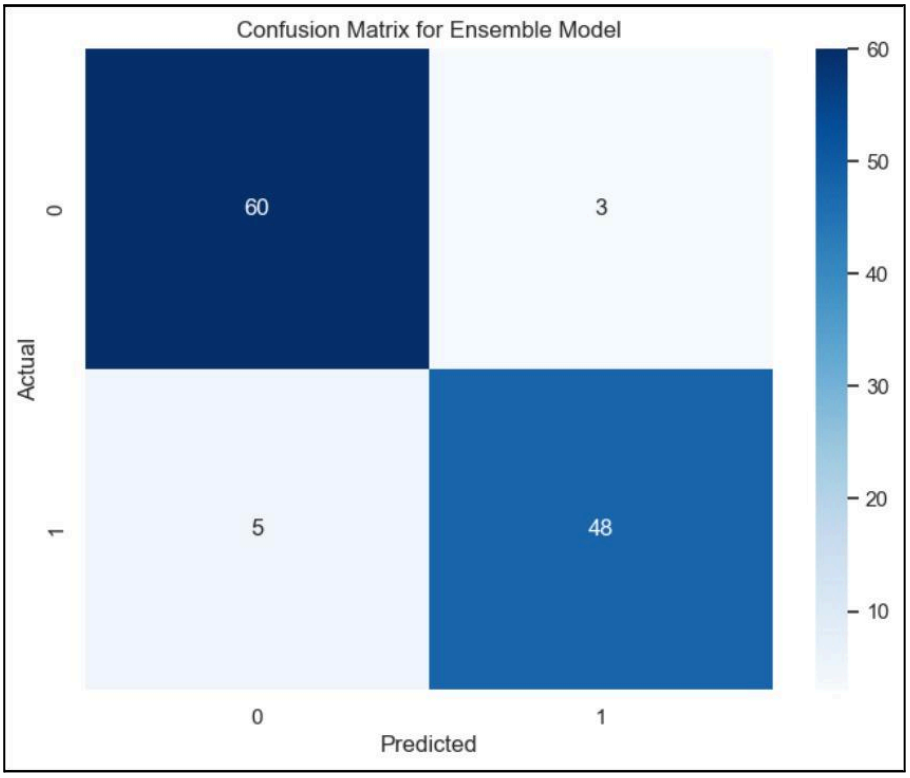


Fig. 13. Confusion Matrix for Ensemble Model

5.6 Limitations

While the evaluation metrics provide valuable insights into model performance, it's essential to acknowledge the limitations of the study. The confusion matrix, for instance, only presents the number of accurate and incorrect predictions for each class, without considering the cost of misclassification or the relative importance of the classes.

In summary, the results and evaluation highlight the effectiveness of the Support Vector Machine, Gradient Boosting, and Random Forest models, as well as the ensemble model, in detecting fake profiles on Instagram. The ensemble model, in particular, demonstrates superior performance, achieving high accuracy and robust predictive power.

CHAPTER 6
CONCLUSION AND FUTURE SCOPE

In this chapter, we provide a summary of the findings from our study on the detection of fraudulent profiles on online social networks (OSNs) and discuss potential avenues for future research and development.

6.1 Conclusion

Through the implementation of machine learning models and a web application, we have demonstrated an effective approach for identifying fraudulent profiles on OSNs. Leveraging features extracted from user profiles, including engagement metrics, textual information, and social interaction indices, our system achieved high accuracy in distinguishing between genuine and fake profiles.

The ensemble voting classifier, combining the predictive power of multiple base models, emerged as the most robust approach, showcasing superior performance compared to individual classifiers. The integration of data preprocessing techniques, feature engineering, and model training facilitated the development of a reliable and scalable solution for profile verification.

The web application developed as part of this study provides users with a user-friendly platform for real-time profile verification, enhancing the security and trustworthiness of online interactions on social media platforms.

6.2 Future Scope

While our study yields promising results, several avenues for future research and development remain to be explored:

1. Enhanced Feature Engineering: Investigating additional features derived from user behavior, network analysis, and content semantics could further improve the accuracy and robustness of the detection system. Features such as posting frequency, content sentiment analysis, and network centrality measures may offer valuable insights into the authenticity of user profiles.

2. Advanced Machine Learning Techniques: Exploring advanced machine learning algorithms, including deep learning architectures such as convolutional neural networks (CNNs) and recurrent neural networks (RNNs), could enhance the system's capability to detect subtle patterns and anomalies in user profiles. Additionally, ensemble methods such as stacking and boosting may further boost classification performance.
3. Real-time Monitoring and Adaptive Learning: Implementing a real-time monitoring system that continuously analyzes user activity and adapts the detection model dynamically can enhance the system's responsiveness to emerging threats and evolving attack strategies. Incorporating feedback mechanisms to update the model based on user interactions and feedback can improve its accuracy over time.
4. Cross-platform Integration: Extending the detection system to encompass multiple social media platforms beyond Instagram, such as Twitter, Facebook, and LinkedIn, would provide comprehensive protection against fraudulent activities across diverse online environments. Developing platform-agnostic features and detection techniques can facilitate seamless integration with different OSNs.
5. User Education and Awareness: Investing in user education and awareness campaigns to educate users about the risks of fraudulent profiles and the importance of vigilance in online interactions can complement technical solutions. Empowering users with knowledge and tools to identify suspicious behaviors and report fraudulent activities can contribute to a safer online community.

In conclusion, our study represents a significant step towards addressing the challenge of fraudulent profiles on online social networks. By combining machine learning techniques with web application development, we have developed a practical solution for profile verification, laying the groundwork for future advancements in online security and trustworthiness.

REFERENCES

[1] Eswara Sai Raja, VS Aditya, and Sachi Nandan Mohanty. "Fake Profile Detection Using Logistic Regression and Gradient Descent Algorithm on Online Social Networks." International Conference on Communication, Computing and Internet of Things (IC310T) (2023). https://publications.eai.eu/index.php/sis/article/view/4342/2676

[2] K. Harish, R. Naveen Kumar, Dr. J. Briso Becky Bell. "Fake Profile Detection Using Machine Learning." IJSRSET 2023 (2023). https://ijsrset.com/paper/8885.pdf

[3] Dr. K. Smita, N. Harika, O Lakshmi Kalyani, T. Kruthika. "FAKE PROFILE IDENTIFICATION IN SOCIAL NETWORK USING MACHINE LEARNING AND NLP" International Conference on Communication, Computing and Internet of Things (IC310T) (2023). https://turcomat.org/index.php/turkbilmat/article/view/14127

[4] Partha Chakraborty, Mahim Musharof Shazan, Mahamudul Nahid, Md. Kaysar Ahmed. "Fake Profile Detection Using Machine Learning Techniques" IEEE International WIE Conference on Electrical and Computer Engineering (WIECON-ECE) (2022). https://www.scirp.org/journal/paperinformation.aspx?paperid=120727

[5] Ajith M, M. Nirmala. "Fake Accounts and Clone Profiles Identification on Social Media Using Machine Learning Algorithms." (2022). https://ijsrset.com/IJSRSET2293158

[6] M.Mamatha, M.Srinivasa Datta, Umme Hani Ansari, Dr. Subhani Shik. "Fake Profile: Identification using Machine Learning Algorithms." (2021). https://www.ijera.com/papers/vol11no7/Ser-3/H1107036065.pdf

[7] Van Der Walt, E., & Eloff, J. (2018). Using Machine Learning to Identify False Identities: Bots vs. Humans. IEEE Access, 6, 6540–6549. https://doi.org/10.1109/ACCESS.2018.2796018

[8] Ferrara, E., & Kudugunta, S. (2018). Bot detection using deep neural networks. Information Sciences, 467, 312–322. https://doi.org/10.1016/j.ins.2018.08.019

[9] Ramalingam, D., & Chinnaiah, V. (2018). Fake Profile Detection Methods in Large-Scale Online Social Networks: A Complete Study. Computers & Electrical Engineering, 65, 165–177. https://doi.org/10.1016/j.compeleceng.2017.05.020

[10] Minoso, Y., Hajdu, G., Lopez, R., Acosta, M., & Elleithy, A. (2019). Artificial Neural Networks are used to spot fake profiles. 2019 IEEE Long Island Systems, Applications, and Technologies Conference (LISAT). https://doi.org/10.1109/LISAT.2019.8817330

[11] Swe, M. M., & Myo, N. N. (2018). Using a black-list, detected fake accounts on Twitter. IEEE/ACIS 17th International Conference on Computer and Information Science (ICIS). https://doi.org/10.1109/ICIS.2018.8466499

[12] Jie, H. J., & Wanda, P. (2020). DeepProfile: Utilizing Dynamic CNN to Detect Fake Profiles in Internet Social Networks. Journal of Information Security and Applications, 52, Article ID 102465. https://doi.org/10.1016/j.jisa.2020.102465

[13] Kodati, S., Reddy, K. P., Mekala, S., Murthy, P. S., & Reddy, P. C. S. (2021). Identification of Fake Accounts on Twitter Using Hybrid SVM Algorithm. E3S Web of Conferences, Article No. 01046. https://doi.org/10.1051/e3sconf/202130901046

[14] Meshram, E. P., Bhambulkar, R., Pokale, P., Kharbikar, K., & Awachat. (2021). Automated Fake Profile Detection Using Machine Learning on Instagram. Journal of Scientific Research in Science and Technology: International, 8, 117–127. https://doi.org/10.32628/IJSRST218330

[15] Chakraborty, P., Muzammel, C. S., Khatun, M., Islam, S. F., & Rahman, S. (2020). Utilizing face recognition, developed an automatic student attendance system. IJEAT, 9, 93–99. https://doi.org/10.35940/ijeat.B4207.029320

[16] Sayeed, S., Sultana, F., Chakraborty, P., & Yousuf, M. A. (2021). Evaluation of Eye-ball Movement and Head Movement Detection Based on Reading. In S. Bhattacharyya, L. Mri, M. Brkljai, J. V. Kureethara, & M. Koeppen (Eds.), Current Developments in Signal and Image Processing (pp. 95–103). Springer. https://doi.org/10.1007/978-981-33-6966-5_10

[17] Forecasting Degree of Visual Focus of Human Attention Using Machine Learning Techniques by P. Chakraborty, M. A. Yousuf, and S. Rahman. In: Proceedings of the International Conference on Trends in Computational and Cognitive Engineering, Springer, Singapore, 683-694. Shamim Kaiser, Bandyopadhyay, Mahmud, and Raym, editors. https://doi.org/10.1007/978-981-33-4673-4 56

[18] Zero-Shot Learning to Identify Item Instances from Unknown Picture Sources. Muzammel, C.S., Chakraborty, P., Akram, M.N., Ahammad,K., and Mohibullah, M. (2020).IJITEE, 9, 988–991. International Journal of Innovative Technology and Exploring Engineering.

https://doi.org/10.35940/ijitee.C8893.029420

[19] Utilizing Template and Hog Feature Matching, Sultana, T. Ahmed, P. Chakraborty, M. Khatun, M. Hasan, and M. S. Uddin published Object Detection in 2020. IJACSA, 11, 233-238. International Journal of Advanced Computer Science and Applications. https://doi.org/10.14569/IJACSA.2020.0110730

[20] Using machine learning techniques, Faruque, M.A., Rahman, S., Chakraborty, P., Choudhury, T., Uh, J.S., and Singh, T.P. (2021) ascertain the polarity of the public's opinions on cricket in Bangladesh.30,
https://doi.org/10.1007/s41324-021-00403-8

[21] Using a machine learning approach based on Twitter data, Sarker, Chakraborty, Sha, Khatun, Hasan, M.R., and Banerjee (2020) developed an improvised technique for data analysis and terrorist attack detection.8, 50–62, Journal of Computing and Communications.
https://doi.org/10.4236/jcc.2020.87005

[22] Identifying Fake Accounts on Social Media, S. Khaled, N. El-Tazi, and H.M. Mokhtar, 2018. Big Data 2018 IEEE International Conference, Seattle, 10–13 December 2018, 3672–3681.
https://doi.org/10.1109/BigData.2018.8621913

[23] Social Networks Fake Profiles Detection Using Machine Learning Algorithms, Y. and Z. Elyusufi, 2019. In: Innovations in Smart Cities Applications