Unmasking Deceptive Profile in Social Network using Machine Learning

V. Navyasri[1], N. Haveela Theresa[2], P. Sai Sathvika[3], I.M.V. Krishna[4]

Department of Information Technology[1,2,3,4]
PVP Siddhartha Institute of Technology[1,2,3,4]

---***---

Abstract - Social media platforms have become a regular aspect of our lives, acting as mediums for networking, information sharing, and communication. However, the proliferation of deceptive profiles poses a significant challenge to the authenticity and integrity of these platforms. Deceptive profiles can be utilized for various malicious purposes such as spreading misinformation, engaging in cyberbullying, or conducting fraudulent schemes. In this research paper, we propose a comprehensive approach for distinguishing between genuine and fake anonymous profiles on social media platforms using machine learning (ML) and natural language processing (NLP). We investigate a range of features and methodologies, including behavioral patterns, sentiment analysis, and profile completeness, also use ML algorithms to effectively discern between genuine and fake profiles. Experimental results obtained from real-world datasets demonstrate the efficacy and scalability of our proposed approach. Moreover, we discuss the implications of fake profiles on social media and propose strategies for mitigation and prevention.

***Key Words*: Deceptive Profile Identification, Machine Learning, Natural Language Processing, Behavioral Patterns, Sentiment Analysis, Profile Completeness.**

1. INTRODUCTION:

The rise of social networking platforms has revolutionized the way individuals interact and communicate online, with millions of users globally engaging in these platforms daily. Online Social Network (OSN) services encompass a wide range of platforms catering to diverse user preferences, from those centered on social interactions like Facebook and Myspace, to information dissemination-centric ones like Twitter and Google Buzz, to platforms integrating social features like Flicker.

However, alongside the immense popularity of social networking comes a host of security concerns and privacy challenges that pose significant obstacles to the seamless operation of OSN platforms. As users engage with social networks, they inevitably share varying degrees of personal information, rendering them susceptible to various forms of cyber threats, including identity theft. Identity theft, characterized by the unauthorized use of personal information for nefarious purposes, has emerged as a pervasive issue plaguing million worldwide. The consequences of identity theft are far-reaching, encompassing financial losses, legal ramifications, reputational damage, and strained interpersonal relationships. Despite the widespread prevalence of identity theft and related cyber threats, many OSN platforms lack robust mechanisms for verifying user accounts and enforcing stringent privacy and security protocols. This deficiency often results in default privacy settings that offer minimal protection to users, thereby rendering OSNs fertile grounds for fraudulent activities and abuse, including identity theft and impersonation attacks. Moreover, the requirement for users to provide accurate personal information during account creation exacerbates the vulnerability of profiles to exploitation by malicious actors.

In the realm of online networks, user profiles can be categorized into static and dynamic data sets. Static data encompasses demographic details and user interests submitted during profile creation, while dynamic data comprises real-time user behavior and network interactions. However, the predominance of static profiles on many social networking platforms complicates efforts to detect fake identities and malicious content effectively. Amidst these challenges, the prevalence of anonymous profiles further exacerbates security concerns and complicates the task of safeguarding user privacy on social networking sites. Anonymous profiles, characterized by the absence of verifiable user information, are often leveraged to perpetrate various forms of misconduct, including online bullying, harassment, and dissemination of misinformation. Despite efforts by platforms like Facebook to implement security measures such as the Facebook Immune System (FIS) to combat spam and phishing attacks, detecting and mitigating the proliferation of fake profiles remains a daunting challenge. In response to these challenges, researchers and practitioners have proposed various approaches to identify and counteract fake identities and malicious activities on online social networks. However, addressing these issues necessitates ongoing research and innovation to develop robust solutions that effectively safeguard user privacy and security while preserving the integrity of social networking platforms.

2. RELATED WORK:

The motivation for this work is in the digital land space, the proliferation of deceptive profiles across online platforms poses a significant challenge to trust and security. Despite existing efforts to detect fraudulent or misleading profiles, traditional methods often fall short in accurately identifying deceptive behaviors. This project aims to address this issue by leveraging machine learning techniques to develop a robust and efficient system capable of unmasking deceptive profiles. By analyzing various features and patterns inherent in user profiles and activities, the proposed solution seeks to enhance the detection and mitigation of deceptive behavior, thereby fostering a safer and more trustworthy online environment

An analysis of Roy, P. K., & Chahar, S. (2020) provides insights into the many approaches and strategies used in the field of false profile identification on social networking platforms. The article offers a comprehensive review of previous research, emphasizing the difficulties and developments in identifying false profiles in online social networks [1]. Joshi, U. D., et al. (2021) explore machine learning algorithms and their applications in identifying deceptive accounts in their discussion of phony social media profile identification. Through the application of machine learning techniques, researchers hope to create strong models that can distinguish between real and deceptive profiles on various social media networks [2]. In Harris, P., et al.'s study from 2021, machine learning is used to identify and categorize deceptive Instagram profiles. Researchers hope to improve the security and reliability of Instagram by using machine learning algorithms to identify which profiles are real and which are false [3].

Another piece of literature is about the Linkenln platform. Shalinda Adikari and Kaushik Dutta's paper (PACIS 2014 Proceedings) explores the topic of identifying phony LinkedIn profiles. Their study clarifies the unique characteristics and actions exhibited by fraudulent accounts on a professional networking site. Through the examination of these traits, the research offers significant understanding into the methods used by those who want to fabricate identities for fraudulent intent. This analysis contributes to our knowledge of dishonest practices on LinkedIn and helps to guide the creation of more robust detection systems that protect the integrity of professional networks [4]. In the endeavor to unveil deceptive profiles within online social networks, a thorough literature review draws upon insights from several notable studies. Smita et al. (2023) contribute significantly to this area by proposing a method that harnesses both natural language processing (NLP) techniques and machine learning for deceptive profile identification. By integrating these advanced technologies, their research offers a comprehensive approach to detecting deceptive behaviors within social networks [5]. Similarly, Vyawahare and Gavaskar (2022) focus on the utilization of profanity and gender identification as key indicators for fake profile recognition. Their study underscores the importance of linguistic and demographic cues in uncovering deceptive practices, thereby enriching the understanding of deceptive behaviors in online environments [6]. Additionally, Kadam and Sharma (2022) present a data mining technique tailored specifically for social media fake profile detection. Their research emphasizes the significance of computational methods in effectively identifying and mitigating deceptive profiles across various social media platforms [7]. Together, these studies provide valuable theoretical foundations and methodological approaches for addressing the pervasive issue of deceptive profiles, offering essential insights for the development of robust detection mechanisms in OSN (online social). networks.

One more comprehensive literature review draws upon insights from several notable studies. Masood et al. (2019) significantly contribute by focusing on spammer detection and fake user identification, offering valuable insights into the identification of deceptive behaviors prevalent in social networks. Their study sheds light on the intricate mechanisms involved in identifying and mitigating deceptive profiles, providing a foundation for understanding the nuances of deceptive practices [8]. Similarly, Kodati et al. (2021) present a hybrid SVM algorithm specifically tailored for detecting fake profiles on Twitter. Their research underscores the importance of employing innovative machine learning approaches to combat the proliferation of deceptive profiles in social media environments [9]. Furthermore, Prabakaran (2022) explores the application of clustering techniques for detecting deceptive profiles in online social media streams. By leveraging data-driven methodologies, Prabakaran's study highlights the effectiveness of clustering algorithms in identifying patterns indicative of deceptive behaviors, thereby enhancing our understanding of deceptive practices in online social networks [10]. Additionally, Sen et al. (2022) propose a novel approach that combines image processing and machine learning for fake profile detection. Their research showcases the potential of multi-modal analysis in improving the accuracy and robustness of detection mechanisms, offering promising avenues for future research in this domain [11]. Together, these studies provide valuable theoretical foundations and methodological approaches for addressing the challenge of deceptive profiles, offering essential insights for the development of robust detection mechanisms in online social networks.

3. PROPOSED SYSTEM:

The proliferation of deceptive profiles across various online platforms has become a pressing concern in today's digital age. These profiles often masquerade as genuine users, engaging in activities that range from spreading misinformation to perpetrating identity theft and financial fraud. Existing methods for detecting fraudulent profiles frequently fall short, as they struggle to keep pace with the evolving tactics employed by malicious actors. There is an urgent need for a more sophisticated and adaptive solution that can accurately differentiate between genuine and fake profiles, thereby enhancing trust and security in the online ecosystem. This paper proposes a novel approach that harnesses the power of ML and NLP techniques such as sentiment analysis, behavioral patterns and profile completeness.et al. to develop a robust system capable of effectively unmasking deceptive profiles. This paper proposes an innovative approach that harnesses the synergy of machine learning and natural language processing techniques to effectively unmask deceptive profiles.

The primary goal of this project is to establish a comprehensive system that leverages ML and NLP techniques to effectively identify and mitigate deceptive profiles on online platforms. By analyzing various features and patterns inherent in user profiles and activities, the proposed solution aims to enhance the detection and mitigation of deceptive behavior, ultimately promoting a more reliable and secure online environment. In this paper, we undertake the task of detecting deceptive profiles by employing various Natural Language Processing techniques for textual data analysis. These techniques include gender detection, sentiment analysis, evaluating friend connections, and assessing profile completeness. By integrating these NLP methodologies, we can train ML model effectively thereby enhancing the accuracy and reliability of deception detection systems. Furthermore, our approach involves employing multiple classification algorithms. We conduct a comprehensive comparison among these algorithms to determine their effectiveness in identifying deceptive profiles accurately.

3.1 METHODLOGY:

To achieve this objective, we begin by assembling a diverse dataset containing both genuine and fake profiles sourced from various online platforms. Next, we extract relevant features from user profiles using NLP techniques, including posting frequency, linguistic patterns, friend connections, and profile completeness. These features serve as the basis for training multiple machine learning algorithms, which are then compared to identify the most effective approach for differentiating between genuine and fake profiles. The integration of ML and NLP techniques facilitates a multi-dimensional analysis, enabling the system to learn the emerging tactics employed by fraudulent actors in creating deceptive profiles. So, below fig 1 shows workflow or Architecture of detecting deceptive profile

3.1.1 Profile Dataset Collection:

In our approach, we collect a diverse dataset containing both genuine and fake profiles sourced from online platforms. We got this collection from Kaggle, which is a website where people share datasets. Having this dataset will help us figure out which profiles might be fake

- **Default Profile Image:** Indicates whether the user is using the default profile image.
- **Geo-Enabled:** Indicates whether the user has enabled location information.
- **Profile Image URL:** Link to the user's profile picture
- **lang**: Language preference of the user.

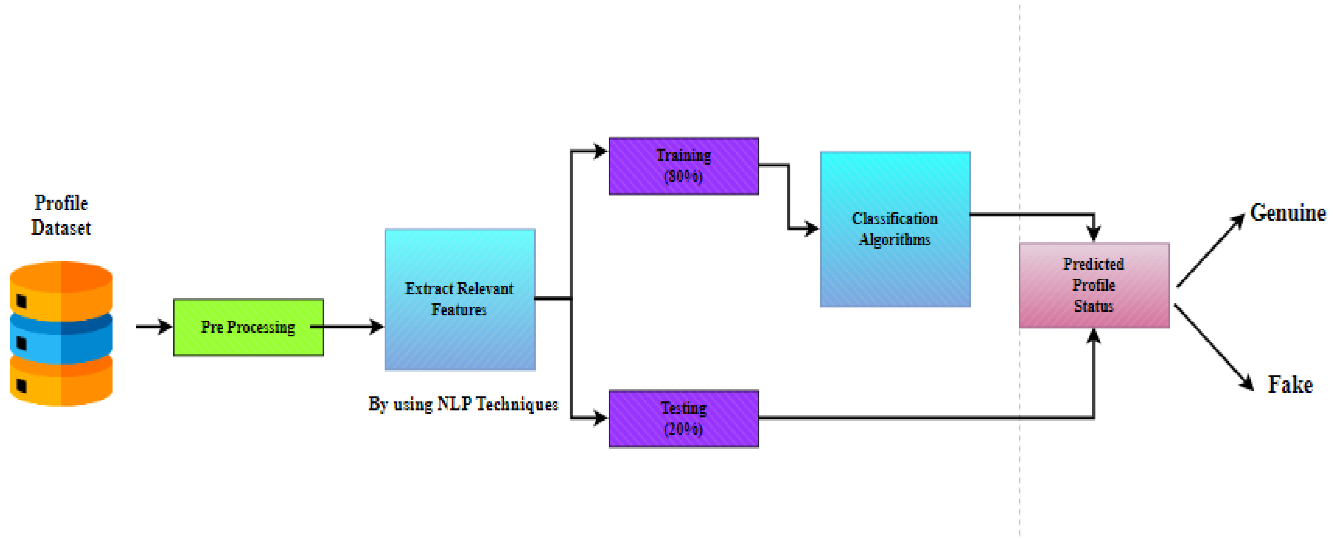

Figure 1: Workflow identification of Unmasking deceptive profiles in the digital era.

- Following are the attributes/ features to Unmask deceptive profile. **id**: A unique number or identifier assigned to each user.
- **name**: The name the user has on their profile.
- **screen_name**: The username or handle of the user.
- **fav_number**: Number of favorites (likes) the user has.
- **Statuses Count:** Total count of posts or updates made by the user.
- **Followers Count:** Number of individuals who follow the user.
- **Friends Count:** Number of individuals whom the user follows.
- **Favorites Count:** Total count of posts or updates liked by the user.
- **Listed Count:** Indicates the number of lists that include the user as a member.
- **URL**: Website link provided in the user's profile (if any).
- **time zone**: Time zone set by the user.
- **location**: Geographical location mentioned by the user.
- **Created At:** Date and time when the user's account was originally created.
- **Banner Image URL:** Link to the user's profile banner image.
- **Background Image Usage:** Indicates whether the user's background image is currently in use.
- **Background Image URL (HTTPS):** HTTPS links to the user's background image.
- **Profile Picture URL (HTTPS):** Secure URL to the user's profile picture.
- **Profile Sidebar Border Color:** Color of the border surrounding the user's profile sidebar.
- **Background Image URL:** Link to the user's profile background image.
- **Default Profile Layout:** Indicates whether the user is using the default profile layout.
- **Background Image Tiling:** Indicates whether the user's background image is tiled.
- **Profile Background Color:** Color of the background in the user's profile.
- **Profile Link Color:** Color of links within the user's profile.
- **utc_offset**: UTC offset set by the user.
- **verified**: Indicates if the user has a verified account.
- **Description**: Description or bio provided by the user.
- **updated**: Date and time when the user's profile was last updated.
- **label**: Label indicating whether the profile is genuine or fake.

3.1.2 Data-Preprocessing:

Data pre-processing is a crucial step in machine learning model creation. It ensures that data is clean and formatted correctly for the model, preventing misleading outcomes. This process involves transforming data, dealing with noise, duplicates, and missing values. Activities in data pre-processing include importing datasets, splitting datasets, and attribute scaling. Pre-processing improves model accuracy by ensuring data is in the right format

3.1.3 Extracted Features:

The main objective of this stage is to Identify Deceptive profiles. But that they may misrepresent or hide certain aspects of the person behind them. So, this is a crucial step in the project where you gather important information from the profiles that can help in identifying deception. Relevant features are specific characteristics or attributes of profiles that might indicate deception.

Here we apply NLP pre-processing techniques to the Extracted Features to transform raw text data into a format that can be easily understood and analyzed by machine learning algorithms. Here are some NLP preprocessing techniques employed in the project

- **Gender-Detection:** Here we are using a technique called gender detection to analyze the names provided in the profiles. This helps in determining whether the name corresponds to a male or female gender, which could be useful in detecting inconsistencies or potential deception. To achieve this task, we imported a library called “gender_guesser.detector” so it first extracts the name from the input, then utilizes a model within the library to predict the gender associated with that name. The predicted gender is then converted into a numerical code for ease of processing. This code can help in tasks like analyzing demographics or personalizing user experiences based on gender.
- **Sentiment Analysis**: Sentiment analysis (Description Attribute) helps in understanding the underlying sentiment or emotion expressed in the text. This can reveal whether the language used in the description is overly positive, negative, or neutral, which might indicate deception or manipulation. It generated some score by analyzing that text. To achieve this, we imported **“SentimentIntensityAnalyzer”** object from NLTK's” **nltk.sentiment.vader”** module. So, it takes piece of text as input and calculates a sentiment score, which indicates the overall positivity or negativity of the text.
- **Profile Completeness:** This feature assesses (default_profile_image , profile_background_tile, time_zone, geo_enabled,URL) how complete or comprehensive the profiles are. So, this NLP methods might be used to analyze the completeness of the profiles, such as checking for missing information or inconsistencies in the provided details.

By applying these techniques, the project aims to unmask deceptive profiles by identifying inconsistencies, patterns of deception, or indicators of manipulation within the extracted features. The goal is to develop a system or model that can automatically flag or identify deceptive profiles based on the analyzed features. So, this process forms the foundation for building a model capable of detecting deceptive profiles in various contexts.

Following are the Extracted Features:

- **statuses_count:** Total number of posts or updates the user has made
- **followers_count:** Number of people following the user.
- **friends_count:** Number of people the user follows.
- **favourites_count:** Total number of posts or updates the user has liked
- **listed_count:** Number of lists the user is a part of.
- **sex_code:** By using gender detection we convert name into sex code.
- **lang_code:** Language preference of the user
- **Description:** Description or bio provided by the user. Here we converted text into sentimental score by applying sentiment analysis
- **profile_com:** By using multiple attributes we calculate average based upon their user preferences

If we fail to extract relevant features, our machine learning model will have a very low accuracy rate. This means that the model won't be able to effectively understand or learn from the data provided to it. The below figure2 illustrate this point, highlighting the correlation between feature extraction quality and the accuracy rates of the multiple models. In essence, the accuracy of our model heavily depends on the quality and relevance of the features we extract from the data**.**

3.1.4 Splitting the dataset:

Before initiating the training phase, it's essential to divide the dataset. Initially, we determine the input and output values, where the input comprises independent variables, and the output consists of the dependent variable, indicating either "Fake" or "Genuine," represented by values 0 and 1, respectively. Subsequently, we partitioned the dataset into two subsets: the training set and the test set. Our approach allocates 80% of the data to the training set and reserves the remaining 20% for the test set

3.1.5 Classification Algorithms:

In the project aimed at unmasking deceptive profiles, various machine learning algorithms have been employed for analysis. These algorithms include Support Vector Machine (SVM), K-Nearest Neighbors (KNN), Decision Tree, and Random Forest. **1. Support Vector Machine (SVM):**

Support Vector Machine (SVM) is a type of supervised learning algorithm primarily utilized for classification tasks. Its core principle involves identifying the hyperplane that effectively separates different classes within the feature space. In the context of unmasking deceptive profiles, SVM can learn to classify profiles as genuine or deceptive based on the features extracted from them. It aims to find the optimal decision boundary that maximizes the margin between the two classes.

K-Nearest Neighbors (KNN):

K-Nearest Neighbors (KNN) is recognized as a straightforward yet robust supervised learning algorithm suitable for both classification and regression tasks. In KNN, when presented with a new data point, its class is determined by the majority class among its nearest neighbors in the feature space. In the context of the project, KNN can be utilized to classify profiles as genuine or deceptive based on the characteristics of similar profiles in the dataset.

Decision Tree:

Decision trees represent hierarchical structures commonly employed in tasks involving classification and regression. They recursively divide the feature space into subsets to make predictions or decisions. Hey recursively partition the feature space into subsets based on the values of features, aiming to minimize impurity or maximize information gain at each node. Decision trees are interpretable and can reveal the most important features for classification. In the project, decision trees can help identify the key features distinguishing genuine profiles from deceptive ones.

Random Forest:

Random Forest is an ensemble learning method that constructs multiple decision trees and combines their predictions to improve accuracy and robustness. Each tree in the forest is built from a random subset of the training data and a random subset of features. Random Forest is effective for handling high-dimensional data and mitigating overfitting. In the context of unmasking deceptive profiles, Random Forest can provide more reliable predictions by aggregating the outputs of multiple decision trees.

Overall, by employing these machine learning algorithms, the project aims to develop a robust and accurate model capable of detecting deceptive profiles based on various extracted features. Each algorithm brings its own strengths and characteristics to the task, contributing to the comprehensive analysis and unmasking of deceptive profiles.

4. Results and Discussion:

Evaluation metrics and Results:

Detecting deceptive profiles in various contexts, such as social media platforms, requires careful consideration of evaluation metrics to assess the effectiveness of detection methods. So, the following are some of the metrics we are used in our paper.

Accuracy: The percentage of deceptive profiles that are accurately identified out of all the profiles is measured by accuracy. It serves as a fundamental metric in assessing the overall performance of detection algorithms. By using formula (1) we get accuracy**.**

Accuracy=Total Amount of Predictions / Number of Correct Predictions. (1)

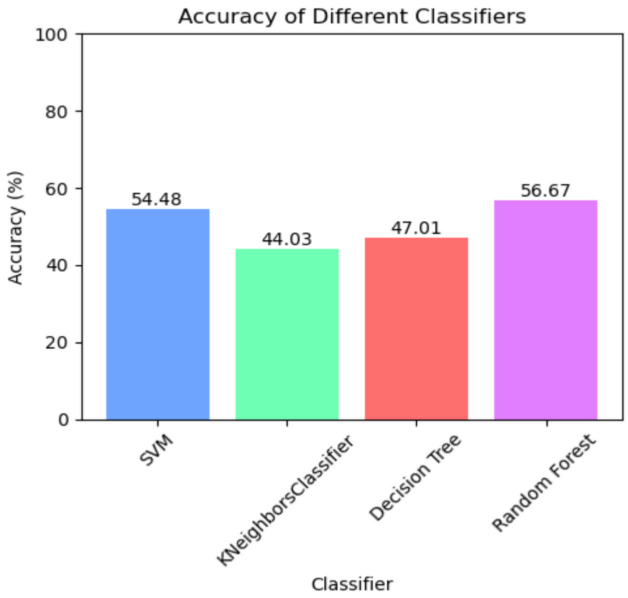

Figure2 represents the accuracy levels of Trained ML models without proper Feature extraction and NLP implementation

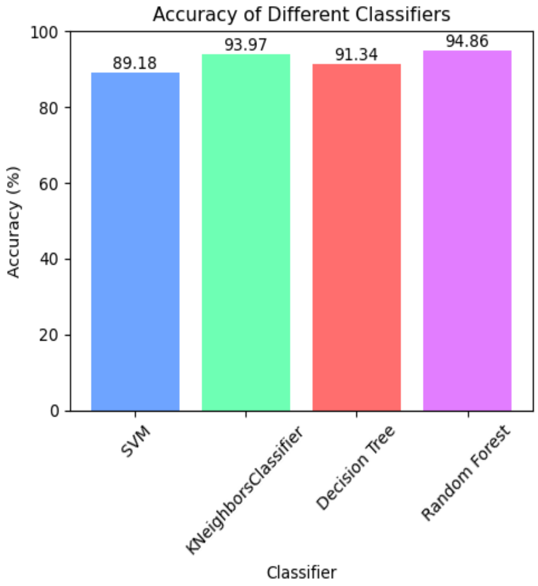

Figure 3 represents final accuracies for trained ML models to identified deceptive profiles among all profiles.

Accuracy measures the overall correctness of the ML model's predictions. According to this metric, the best model would be considered the best with the highest accuracy. According to the information provided, the Random Forest model achieves the highest accuracy among the models listed, with an accuracy of 99%.

Confusion Matrix: A confusion matrix presents a comprehensive breakdown of the performance of the detection system, including true positives, true negatives, false positives, and false negatives. It offers insights into the specific types of errors made by the algorithm

Table 1: Confusion matrix

	Predicted **Negative**	**Predicted** **Positive**
Actual Negative	True Negatives (TN)	False Positives (FP)
Actual Positive	True Negatives (FN)	True Positives (TP)

Table 2: Classification of deceptive profiles on Facebook dataset using multiple algorithms such as SVM, KNN, Decision tree, Random Forest

Model	**TN**	**FP**	**FN**	**TP**
SVM	232	36	24	272
KNN	260	26	8	270
Decision tree	240	28	24	272
Random Forest	268	0	29	267

- True Positives (TP): The number of correctly identified deceptive profiles.
- True Negatives (TN): The number of correctly identified non-deceptive profiles.
- False Positives (FP): The number of non-deceptive profiles incorrectly classified as deceptive.
- False Negatives (FN): The number of deceptive profiles incorrectly classified as non-deceptive.

Precision and Recall:

- Precision calculates the percentage of deceptive profiles among all profiles classified as deceptive.
- Recall measures the proportion of correctly identified deceptive profiles among all truly deceptive profiles in the dataset.
- The F1-score represents the harmonic mean of recall and precision, is often used to balance these two metrics.

Model	**precision**	**Recall**	**F1-Score**
SVM	0.90	0.85	0.87
KNN	0.974	0.914	0.943
Decision tree	0.91	0.90	0.90
Random forest	1.0	0.97	0.989

Table 3: Comparative results of Precision, recall, F1-Score for Different algorithms

Here Precision indicates how reliable the positive predictions of the model are. In this metric, the Random Forest model also performs the best, with a precision of 100%. And recall tells us how effectively the model captures all positive instances. Again, the Random Forest model achieves the highest recall at 97%. Finally, F1-score is useful metric when there's an uneven class distribution. Once more, the Random Forest model outperforms others with an F1-score of 98.9%.

Considering all evaluation metrics provided, the Random Forest model consistently performs the best across accuracy, precision, recall, and F1-score. It achieves the highest values in each metric, indicating its superior performance in identifying deceptive profiles on the Facebook dataset compared to SVM, KNN, and Decision Tree models.

5. CONCLUSION:

The online version of the volume will be available in LNCS Online. Members of institutes subscribing to the Lecture Notes in Computer Science series have access to all the pdfs of all the online publications. Non-subscribers can only read as far as the abstracts. If they try to go beyond this point, they are automatically asked whether they would like to order the pdf and are given instructions as to how to do so.In summary, the proposed system integrates Machine Learning (ML) and Natural Language Processing (NLP) techniques to effectively identify and mitigate deceptive profiles on online platforms. It emphasizes the critical role of feature extraction quality, where proper NLP preprocessing techniques transform raw text data into a format suitable for ML analysis, thereby enhancing accuracy rates. Various evaluation metrics such as accuracy, confusion matrix, F1-Score, precision and recall are utilized to assess the effectiveness of detection methods, providing insights into ML algorithm performance. Among the tested ML algorithms, The Random Forest model is popular as the most effective in distinguishing between genuine and deceptive profiles, highlighting its superiority in this domain. However, challenges such as data quality, scalability, and privacy concerns need to be addressed for future improvements. The proposed system adopts a comprehensive approach to deception detection, considering diverse attributes and behaviors in user profiles. By combining advanced ML and NLP techniques, the system can adapt to evolving tactics used by malicious actors, ultimately fostering a safer and more trustworthy online environment.

5.1 Limitation of Current Research

Data Quality and Diversity: The effectiveness of ML models greatly depends on the quality and variety of the dataset used for training. Limited access to diverse and high-quality data may lead to biased or less accurate models, especially in detecting deceptive profiles that exhibit subtle variations across different platforms and demographics

Scalability and Efficiency: Deploying ML-based deception systems at scale requires efficient algorithms and infrastructure to manage massive data sets in real time. Ensuring scalability and efficiency while maintaining detection accuracy is a significant challenge, particularly for resource-constrained platforms.

Privacy Concerns: The use of NLP techniques to analyze user-generated text raises privacy concerns, as it involves processing sensitive personal information. Ensuring user privacy while maintaining the effectiveness of deception detection algorithms is a delicate balance that requires careful consideration.

5.2 Future Scope

The proposed system utilizes machine learning (ML) and natural language processing (NLP) techniques to effectively identify and mitigate deceptive profiles on online platforms. This approach enhances the detection and mitigation of deceptive behavior, promoting an online environment that is more reliable and safer. Future advancements should focus on enhancing data quality and diversity. Additionally, addressing privacy concerns associated with NLP techniques is paramount, and exploring privacy-preserving methods like differential privacy or federated learning can ensure user privacy while maintaining the effectiveness of deception detection algorithms. Integration of advanced ML techniques such as deep learning and ensemble methods can further improve accuracy and robustness, while developing real-time monitoring capabilities and adaptive systems using streaming analytics and reinforcement learning can enable continuous adaptation to emerging threats effectively.

ACKNOWLEDGEMENT

We extend our heartfelt gratitude to the authors of the research articles whose work laid the foundation for our project. Their insights and expertise not only provided us with invaluable advice but also inspired us to explore new avenues in our research. We also express our sincere appreciation to our teachers for their continuous support and encouragement throughout this endeavor. Their guidance and dedication have been instrumental in shaping our approach and refining our strategies.

REFERENCES

1. Roy, P. K., & Chahar, S. (2020). Fake profile detection on social networking websites: a comprehensive review. IEEE Transactions on Artificial Intelligence, 1(3), 271-285.
2. Joshi, U. D., Vanshika, Singh, A. P., Pahuja, T. R., Naval, S., & Singal, G. (2021). Fake social media profile detection. Machine Learning Algorithms and Applications, 193-209.
3. 3. Harris, P., Gojal, J., Chitra, R., & Anithra, S. (2021, October). Fake Instagram profile identification and classification using machine learning. In 2021 2nd Global Conference for Advancement in Technology (GCAT) (pp. 1-5). IEEE
4. 4. Shalinda Adikari and Kaushik Dutta, Identifying Fake Profiles in LinkedIn, PACIS 2014 Proceedings, AISeL Shalinda Adikari and Kaushik Dutta, Identifying Fake Profiles in LinkedIn, PACIS 2014 Proceedings, AISeL.
5. Smita, K., Harika, N., Advaitha, N., Kalyani, O. L., & Kruthika, T. (2023). FAKE PROFILE IDENTIFICATION IN SOCIAL NETWORK USING MACHINE LEARNING AND NLP. Turkish Journal of Computer and Mathematics Education (TURCOMAT), 14(03), 689-699
6. Vyawahare, M., & Govilkar, S. (2022). Fake profile recognition using profanity and gender identification on online social networks. Social Network Analysis and Mining, 12(1), 170
7. Kadam, N., & Sharma, S. K. (2022). Social media fake profile detection using data mining technique. Journal of Advances in Information Technology Vol, 13(5)
8. Masood, F., Almogren, A., Abbas, A., Khattak, H. A., Din, I. U., Guizani, M., & Zuair, M. (2019). Spammer detection and fake user identification on social networks. IEEE Access, 7, 68140-68152.
9. Kodati, S., Reddy, K. P., Mekala, S., Murthy, P. S., & Reddy, P. C. S. (2021). Detection of Fake Profiles on Twitter Using Hybrid SVM Algorithm. In E3S Web of Conferences (Vol. 309, p. 01046). EDP Sciences
10. Prabakaran, M. (2022). DETECTING FAKE PROFILES IN ONLINE WITH CLUSTERING TECHNIQUES IN SOCIAL MEDIA STREAMS. Turkish Journal of Computer and Mathematics Education (TURCOMAT), 13(03), 1531-1541.
11. Sen, S., Islam, M. I., Azim, S. S., & Hossain, M. I. (2022, September). Fake Profile Detection Using Image Processing and Machine Learning. In International Conference on Machine Intelligence and Emerging Technologies (pp. 294-308). Cham: Springer Nature Switzerland.