International Journal of Scientific Research & Engineering Trends
Volume 10, Issue 4, July-Aug-2024, ISSN (Online): 2395-566X

Fake Social Media Profile Detection: A Hybrid Approach Integrating Machine Learning and Deep Learning Techniques

Anila S, Meenakshi Mohan, Mariya Jacob, Najiya Nasrin
Adi Shankara Institute of Engineering and Technology Ernakulam, India

Abstract- In the contemporary era of rapid information dissemination through social platforms, the proliferation of fake content undermines the trust and integrity of online communities. Existing detection algorithms exhibit limitations in terms of accuracy and adaptability, necessitating the creation of an innovative hybrid model. Our goal is to integrate the strengths of traditional machine learning approaches, such as k- Nearest Neighbors or Support Vector Machines, with the power of deep learning methods. By combining these techniques, we aim to enhance the accuracy and efficiency of fake profile detection beyond current state-of-the-art methods, providing a robust and effective solution for distinguishing between genuine and deceptive profiles in the dynamic landscape of social media.

Index Terms- machine learning, deep learning, fusion, social media, fake profile, detection

I. INTRODUCTION

In the current landscape of social media, the rapid dissemination of information serves as both a blessing and a curse. While it facilitates the seamless exchange of ideas, it also fosters the proliferation of fake content, often propagated through deceptive social media profiles. The pervasive presence of fraudulent profiles erodes the trustworthiness of online communities, necessitating the development of robust mechanisms for their detection and mitigation. Despite commendable efforts, current algorithms employed for fake social media profile detection exhibit notable limitations in accuracy and adaptability.

Recognizing the urgent need for a more sophisticated solution, our project endeavors to introduce a groundbreaking hybrid model. This model integrates the strengths of traditional machine learning algorithms, such as k-Nearest Neighbors (KNN) and Support Vector Machines (SVM), with the powerful capabilities of deep learning methods. By amalgamating these distinct approaches, we aim to surpass the capabilities of existing state-of-the-art methods in terms of accuracy and efficiency. The primary objective of the project is to develop a comprehensive and advanced framework that transcends the limitations of individual methodologies. Leveraging the synergy between traditional machine learning and deep learning, our hybrid model strives to overcome the inherent shortcomings of each approach. We aspire to convert the weaknesses of one methodology into strengths for the other, thereby constructing a more adaptable and robust system for detecting fake social media profiles.

II. RELATED WORKS

A system addressing the pressing issue of fake profiles in social media and online communities [1] was proposed, recognizing the threat they posed to trust and security. It utilized techniques such as profile completeness analysis, image analysis, linguistic analysis, social network analysis, behavioral analysis, and machine learning models to automatically detect deceptive profiles. Technologies such as soft computing, machine learning, CAPTCHA tests, face detection, image analysis, linguistic analysis, social network analysis, and behavioral analysis contributed to its robustness. The system leveraged algorithms like Eigenfaces, LBP, Adaboost, Support Vector Machine (SVM), and Naive Bayes for tasks ranging from face recognition to texture classification and overall pattern recognition. By incorporating diverse technologies and algorithms, it aimed to stay ahead of evolving tactics employed by malicious actors. Advantages of the system included safeguarding the integrity and trustworthiness of online platforms, enhancing security by identifying and mitigating deceptive activities, protecting users from scams and privacy breaches, and improving the overall user experience by reducing the presence of fake accounts and spam. However, the system faced challenges, including the potential for false positives and false negatives in the detection process, raising concerns about misidentification or missed identification of fake profiles. Ethical considerations related to the collection and analysis of user data, along with the continuous adaptation of malicious actors, presented additional hurdles in the quest to stay ahead in detecting fake profiles. Despite these challenges, the system represented a significant step towards fortifying online platforms against the pervasive threat of deceptive profiles.

Automatic detection techniques for identifying unreliable users on social media are explored in [2]. The study specifically focuses on tackling the spread of fake news. Deep learning methods were employed, utilizing a neural network architecture that combined long short-term memory (LSTM) and convolutional neural network (CNN). The study also included the use of baseline classifiers such as linear support vector machine (SVC), support vector machine optimized by stochastic gradient descent (SVM-SGD), and k-nearest neighbor (KNN) for comparison. Python, TensorFlow, and Keras were among the technologies utilized, with features extracted from public data and metadata to analyze user social context. The system aimed to predict both news and user profile reliability, employing labeled datasets for offline analysis and real user evaluations for online analysis. The integration of deep learning with social context analysis enabled effective detection of fake news and unreliable users. Experimental results demonstrated high accuracy in classifying news content and predicting Twitter profile reliability. However, challenges arose in online analysis, where real users assessed Twitter profiles, potentially impacting classification accuracy. Despite these challenges, the research offered valuable insights for developing more effective tools to combat misinformation on social media.

In the task of detecting fake profiles in social media, the utilization of machine learning algorithms is pivotal, as discussed in [3]. Three supervised machine learning algorithms—Random Forest, Decision Tree (J48), and Naïve Bayes—are employed to identify the most efficient approach. The research underscores the significance of preemptively identifying fake profiles before users are notified to mitigate potential harm. Implemented with the Jupyter tool in Python3, the model's architecture involves data preprocessing, attribute reduction, and the utilization of an ensemble classifier to enhance prediction accuracy. The three-tiered process encompasses profile feature extraction, classification using the selected algorithms, and a subsequent accuracy rate comparison. Notably, the Random Forest algorithm emerges as the most effective in distinguishing between legitimate and fake profiles. Striving to combat advanced persistent threats and malicious activities linked with fake profiles, the research offers a forward-looking perspective, suggesting future research directions such as incorporating ontology engineering for semantic analysis and extending similar approaches to domains with limited information. Key advantages of the proposed approach include its focus on identifying fake profiles before user notification, preventing potential harm caused by these profiles, and demonstrating high accuracy even with limited profile data. However, the document does not explicitly mention any identified disadvantages or limitations associated with the proposed approach or the employed algorithms.

The critical issue of detecting fake profiles on social media platforms, emphasizing the escalating concerns related to misinformation, fraudulent activities, and identity deception associated with these profiles is extensively discussed in [4].

The study emphasizes the importance of early detection using machine learning techniques to improve the safety and trustworthiness of social media platforms. Challenges and drawbacks linked to the expanding use of online social networks, including identity theft and privacy breaches, are also discussed. Various technologies and algorithms, including Random Forest, XGBoost, ADB, GBM, LR, SVM, Neural Networks, KNN, Naïve Bayes, and JRip are explored for accurately identifying fake accounts. The study suggests combining Decision Tree, Naïve Bayes, and SVM algorithms to enhance accuracy, with Gradient Boosting achieving up to 95% accuracy in detecting fraudulent accounts. Features such as profile picture, name, location, age, gender, and user behavior metrics are used for classification. The study acknowledges that the model's performance hinges on the quality of the dataset, feature selection, and algorithm choice. Although specific architecture details are not provided, the methodology involves training a model using known fake and genuine profiles, with machine learning algorithms extracting features and classifying profiles accordingly. Advantages include early detection, accurate identification, and adaptability to new techniques used by fake profile creators. The integration of NLP techniques and the potential analysis of user engagement behaviors contribute to enhanced accuracy in fake profile detection. However, the study acknowledges challenges, including the dependence on dataset quality, algorithm choice, and the need for periodic updates to combat evolving fake profile creation techniques. The study recognizes the limitations in current approaches, leaving room for improvement in the ongoing quest to effectively detect and mitigate the impact of fake profiles on social media platforms.

The research paper extensively addresses the critical issue of spam detection and identification of fake users on social networking sites, specifically focusing on Twitter in [5]. Recognizing the increasing prevalence of spam and its detrimental effects, the document provides a thorough review of techniques for detecting spammers, categorizing them based on their ability to identify various types of spam. The taxonomy includes fake content, URL-based spam, spam in trending topics, and fake users. The study employs various technologies and algorithms, such as clustering, decision trees, Naïve Bayes, machine learning algorithms, Hidden Markov Model (HMM), matrix factorization, and Whiteprint. Feature engineering is highlighted, incorporating user features, content features, graph features, structure features, and time features for a comprehensive analysis. The paper emphasizes the importance of feature selection for improving accuracy and efficiency. A real-time architecture is presented, involving data extraction, a filtering system, data analysis for spammer detection, an alert subsystem, and feedback analysis. The proposed methods offer advantages such as effectiveness in real-time spam detection, feature selection capabilities, a comprehensive approach covering various types of spam, and a detailed comparison of techniques. However, the paper acknowledges challenges and potential disadvantages, including a possible performance trade-off between detection accuracy and processing speed, challenges associated with certain features, the cost of spam removal, a scarcity of publicly available datasets, and the need to consider limitations of machine learning algorithms. In conclusion, the research provides valuable insights into Twitter spam detection, serving as a resource for researchers and addressing the complex challenges of identifying and combating spam and fake users on social media platforms.

The research paper addresses the critical issue of detecting fake accounts on Twitter, highlighting the challenges posed by cybercriminals who exploit social media for spreading false news or stealing user accounts [6]. The proposed mechanism utilizes a stack ensemble system with preprocessing techniques, feature extraction using the Spearman correlation coefficient and chi-square test, and supervised machine learning algorithms. The stack ensemble system, incorporating Random Forest, Support Vector Machine, Naïve Bayes, and Logistic Regression, achieves a remarkable data accuracy of 99%, surpassing the performance of individual algorithms used separately. Advantages of the proposed mechanism include its high accuracy, feature extraction and selection techniques, data cleaning procedures, and the utilization of a stack ensemble system for classification. The mechanism is scalable and specifically designed for Twitter data. However, potential disadvantages include the loss of data during collection errors, the removal of columns with a high percentage of empty fields impacting information loss, dependence on the selection and tuning of individual algorithms in the ensemble, platform specificity to Twitter, and a lack of information on computational resource requirements or potential time constraints. The research does not explicitly discuss limitations or challenges associated with the feature extraction and selection methods employed, urging further consideration for broader applicability.

III. METHODOLOGY

The objective is to effectively detect fake social media profiles by integrating machine learning and deep learning techniques. The process of developing the fake social media profile detection system involves several stages, each meticulously designed to ensure the robustness, accuracy, and reliability of the hybrid model. It consists of data collection, preprocessing, model development, evaluation, validation, and user interaction.

A diverse dataset is collected from Kaggle, comprising genuine and deceptive profiles from Instagram. Preprocessing ensures data quality through steps like removing duplicates, handling missing values, and normalizing features.

Machine learning models, including Support Vector Machine (SVM) and K-Nearest Neighbors (KNN), classify profiles as fake or genuine, with the selection of the more accurate model based on performance. SVM segregates data using hyperplanes, while KNN relies on nearest neighbors for classification.

Deep learning models, such as Multi-Layer Perceptron (MLP) and Long Short-Term Memory (LSTM), capture patterns in textual and sequential data, with the more accurate one chosen. MLP identifies complex patterns via interconnected neurons, while LSTM learns dependencies over time.

Performance evaluation, using metrics like accuracy and precision, ensures model robustness through cross-validation. By leveraging the complementary strengths of these models, the fusion model achieves improved accuracy in detecting fake social media profiles. The weighted fusion technique aims to combine the predictions of traditional machine learning algorithms and deep learning models to leverage their complementary strengths and improve overall detection performance.

The developed hybrid model undergoes comprehensive evaluation using a range of performance metrics to assess its efficacy in detecting fake social media profiles. These metrics include accuracy, precision, recall, F1-score, area under the ROC curve (AUC-ROC), and Mean Average Precision (mAP). The proposed methodology aims to develop an advanced detection system for fake social media profiles by leveraging the synergies between traditional machine learning and deep learning techniques.

IV. EXPERIMENTAL SETUP

1. Data Collection and Preprocessing
Dataset Selection
The choice of dataset is critical in ensuring the reliability and relevance of the experimental findings. The 'instagram-fake-spammer-genuine-accounts' dataset from Kaggle is selected due to its comprehensive nature and suitability for social media profile analysis tasks. This dataset contains a diverse range of features that are essential for detecting fake profiles, including:

- **Profile Picture:** A binary indicator of whether an account has a profile picture or not.

- **Numerical Characters to Username Length Ratio:** A continuous feature representing the proportion of numerical characters in an account's username.
- **Full Name Word Count:** A continuous feature indicating the total number of words in the person's full name.
- **Numerical Characters to Full Name Length Ratio:** A continuous feature indicating the ratio of numerical characters to the total length of the person's full name.
- **Name Matches Username:** A binary feature indicating whether the person's name matches their username.
- **Description Length:** Length of the profile description, likely the bio.
- **External URL in Bio:** A binary feature indicating whether a profile has a link to an external website in its bio.
- **Private Profile:** A binary feature indicating whether the profile is restricted to non-followers.
- **Number of Posts:** A continuous feature representing the total number of posts on the profile.
- **Number of Followers:** A continuous feature representing the total number of followers for each account.
- **Number of Follows:** A continuous feature representing the total number of accounts that the user is following.
- **Fake Account:** Target variable indicating whether an account is fake or not.

The inclusion of these features enables a holistic analysis of Instagram profiles, capturing both textual and visual cues that are indicative of fake or genuine accounts.

In here, we assess the correlation between features, which provides insights into their similarities. The correlation values range from -1 to 1, where -1 indicates a low level of similarity between two features, and 1 indicates a high level of similarity.

It shows the 12 features of the dataset with numerical values and descriptive labels, when the value comes to 1 it represent the similarity of that corresponding feature. It begins with the feature"profile pic" and contains a sequence of numerical values (0.3, 0.36, 0.23, 0.15, 0.17) followed by other sections such as "nums/length username," "fullname words," "nums/length fullname," "name==username," "description length," "external URL," "private," "#posts," "#followers," "#follows," and "fake." Each section feature accompanied by other specific features for detecting the parameters or attributes in a profile.

The numerical values, such as 1, 0.3, 0.22, 0.36, 0.28, 0.23, 0.15, 0.17, and 0.16, suggest some form of scoring or evaluation related to the corresponding features. These values are represent a rating or a measure of compliance with certain standards or expectations for a profile. Additionally, the presence of terms like "private" and "fake" implies that the assessment may involve aspects of privacy settings and profile authenticity.

This graphical structure and content suggest a systematic approach to evaluating or analyzing social media profiles, likely for the purpose of verifying their authenticity, completeness, or adherence to specific guidelines. The inclusion of various metrics, such as profile picture quality, username characteristics, post and follower counts, and the presence of an external URL, indicates a comprehensive assessment of different profile elements. The "fake" label at the end may denote a determination of a profile's authenticity or genuineness based on the established criteria. Overall, it is used to outline a methodical process for evaluating and categorizing social media profiles, potentially aimed at identifying genuine, complete, and compliant profiles.

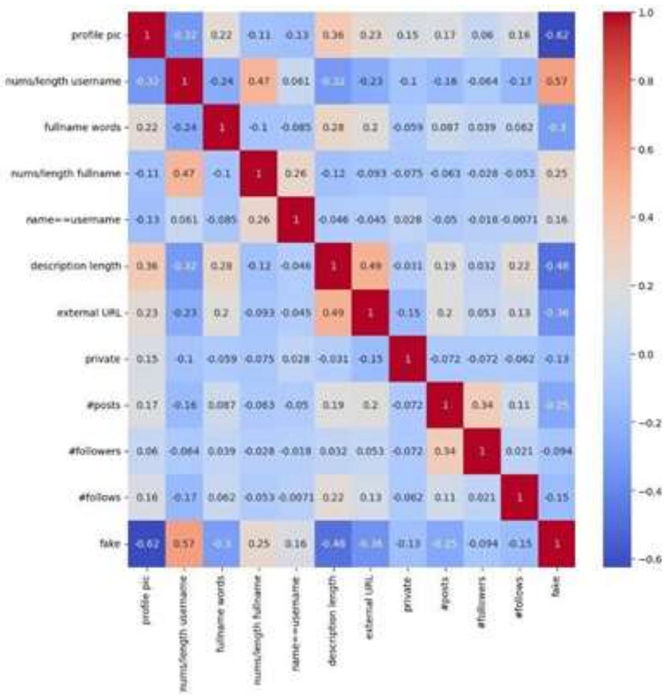

Figure 1: Correlation Matrix of Profile Features and Fake Account Indicator

Feature Engineering

Upon acquiring the dataset, a thorough feature engineering process is conducted to extract meaningful information for analysis. The extracted features include but are not limited to: Profile Picture Presence: A binary feature indicating whether the profile has a profile picture or not, which can be indicative of authenticity.

Username Characteristics

Features such as the length of the username and its alphanumeric composition (e.g., presence of special

International Journal of Scientific Research & Engineering Trends
Volume 10, Issue 4, July-Aug-2024, ISSN (Online): 2395-566X

characters, numbers) are extracted to identify patterns associated with fake profiles.

Full Name Word Count
The number of words in the full name field is considered, as excessively long or nonsensical names may be indicative of fake profiles.

Bio Description Length
The length of the bio description is analyzed as shorter or vague descriptions may raise suspicion.

Post Frequency and Engagement Metrics
Metrics such as the number of posts, likes, comments, and follower-to-following ratio are included to assess the activity and engagement level of the profile, which can be indicative of genuine or fake behavior.

Data Preprocessing
Before training the machine learning and deep learning models, several preprocessing steps are performed to ensure data quality and enhance model performance:

Removing Duplicate Entries
Duplicate entries, if present, are removed to avoid bias in the analysis and ensure each profile is represented uniquely.

Handling Missing Values
Missing values in the dataset are handled through imputation or deletion based on the feature's importance. For numerical features, mean or median imputation may be used, while categorical features may be imputed with the mode or treated as a separate category.

Normalizing Numerical Features
Numerical features are normalized to a common scale (e.g., using Min-Max scaling or standardization) to prevent features with larger magnitudes from dominating the model training process.

Encoding Categorical Features
Categorical features such as profile type (fake, genuine) or account verification status are encoded using techniques like one-hot encoding or label encoding to convert them into numerical representations that can be processed by machine learning algorithms.

2. Machine Learning Models
Support Vector Machine (SVM)
SVM is a powerful supervised learning algorithm used for classification tasks. In this experiment, SVM is trained on the preprocessed dataset to classify profiles as fake or genuine based on the extracted features.

The SVM hyperparameters such as the kernel type and regularization parameter are tuned using techniques like grid search or cross-validation to optimize performance.

Accuracy: SVM classifier achieves an accuracy of 53.43%.

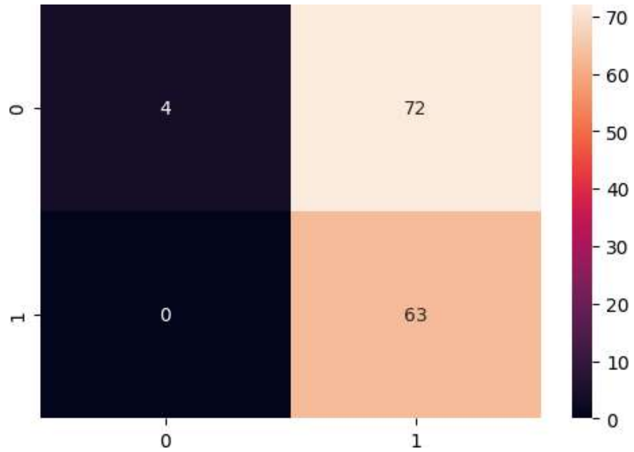

Figure 2: Confusion Matrix of SVM Classifier for Fake Account Detection

The confusion matrix for the SVM classifier reveals that it primarily detects non-fake profiles, correctly identifying 63 instances, while its performance in detecting fake profiles is limited, accurately identifying only 4 instances of fakes.

K-Nearest Neighbors (KNN)
KNN is a non-parametric and instance-based learning algorithm known for its simplicity and effectiveness in classification tasks.

The KNN algorithm classifies profiles by comparing them to the k-nearest neighbors in the feature space. The optimal value of k is determined through experimentation and validation.

Accuracy: KNN classifier achieves an accuracy of 84.12%. Upon examining the results of the KNN classifier, we opted to tune the hyperparameter for the number of neighbors, given our binary classification task (fake and non-fake classes).

This adjustment resulted in an improved accuracy level and a reduction in false positives, as evidenced by the diagonal values in the confusion matrix. Specifically, the correct identification of fake profiles increased to 73 instances, while the correct identification of non-fake profiles increased to 43 instances.

Performance Evaluation
The performance of SVM and KNN models is evaluated using metrics such as accuracy, precision, recall, and F1 score on the testing dataset.

International Journal of Scientific Research & Engineering Trends
Volume 10, Issue 4, July-Aug-2024, ISSN (Online): 2395-566X

The confusion matrix is also analyzed to understand the model's behavior in classifying fake and genuine profiles. From that chose KNN as the accurate one.

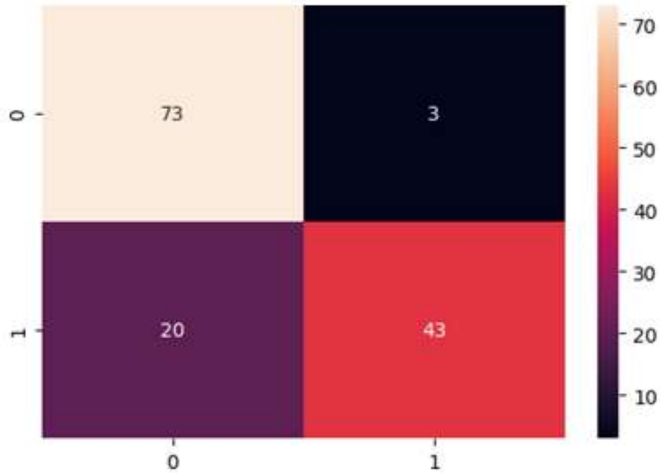

Figure 3: Confusion Matrix of KNN Classifier for Fake Account Detection

3. Deep Learning Models
Multi-Layer Perceptron (MLP)

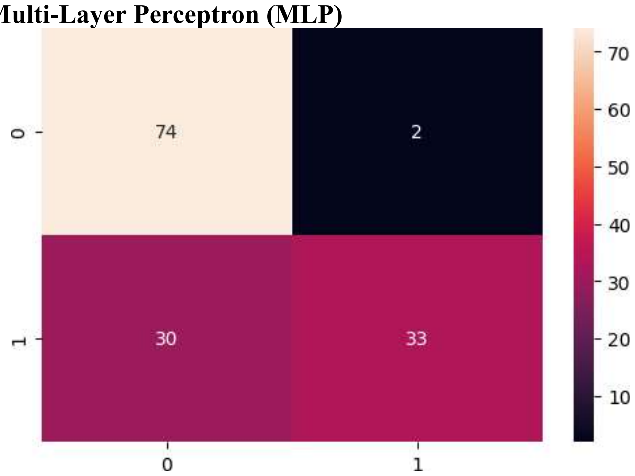

Figure 4: Confusion Matrix of MLP Classifier for Fake Account Detection

MLP is a type of feedforward neural network commonly used for complex pattern recognition tasks. In this experiment, an MLP classifier is trained using the preprocessed data to capture intricate relationships between features and classify profiles accurately.

Accuracy
MLP classifier achieves an accuracy of 79.06%. The confusion matrix plotted for the MLP classifier reveals 74 true positives for the fake class and 33 true positives for the non-fake class. Additionally, it shows 30 false positives for the fake class and 2 false positives for the non-fake class.

Long Short-Term Memory (LSTM)
LSTM is a specialized neural network architecture designed for sequential data analysis, making it suitable for processing text and time-series data. In this context, LSTM is employed to analyze textual features such as bio descriptions and post content to detect patterns indicative of fake profiles.

Model Training and Optimization
The MLP and LSTM models undergo training using techniques like stochastic gradient descent (SGD) or Adam optimizer with appropriate learning rates and regularization techniques to prevent overfitting. Hyperparameter tuning is performed to achieve optimal model performance.

4. Fusion Model
To enhance the accuracy and robustness of fake profile detection, a fusion model is proposed. This fusion model combines the strengths of the most accurate machine learning model (KNN) with deep learning models (MLP and LSTM). The fusion process may involve ensemble techniques such as stacking or blending to leverage the complementary nature of different algorithms.

V. RESULTS AND DISCUSSION

In exploring the effectiveness of various classifiers in identifying fake profiles on social media platforms, it becomes evident that each model possesses distinct strengths and weaknesses.

While the Support Vector Machine (SVM) classifier demonstrated moderate accuracy, its primary strength lay in identifying non-fake profiles, suggesting limitations in accurately detecting fake profiles. In contrast, the K-Nearest Neighbors (KNN) classifier outperformed SVM, achieving an impressive accuracy of 84.12% after hyperparameter tuning, highlighting its effectiveness in this task. Among the deep learning models, the Multi-Layer Perceptron (MLP) showed competitive performance, albeit with a higher number of false positives for fake profiles, while Long Short-Term Memory (LSTM) networks exhibited promise in analyzing textual features but may require further optimization. The proposed fusion model, combining KNN with MLP and LSTM, presents a promising approach to enhance detection accuracy and robustness by leveraging the complementary strengths of different models.

Despite these advancements, the study acknowledges limitations such as the need for larger and more diverse datasets and ethical considerations. Overall, the research contributes valuable insights into the detection of fake profiles, emphasizing the significance of KNN and the potential of fusion models to improve security and trustworthiness in online communities.

VI. CHALLENGES AND SOLUTION

One of the primary challenges encountered in the experimental setup for detecting fake profiles on social media platforms is the handling of imbalanced datasets, where the number of genuine profiles significantly outweighs the number of fake profiles. This imbalance can lead to biased model performance, with classifiers favoring the majority class and exhibiting poor sensitivity towards the minority class. To address this challenge, techniques such as oversampling of the minority class, undersampling of the majority class, or using advanced sampling methods like SMOTE (Synthetic Minority Over-sampling Technique) can be employed to rebalance the dataset and improve model performance. Additionally, another challenge lies in feature selection and engineering, where determining which features are most informative for distinguishing between genuine and fake profiles can be complex. Solutions to this challenge involve conducting thorough exploratory data analysis (EDA) to identify relevant features, leveraging domain knowledge, and employing feature selection algorithms such as recursive feature elimination (RFE) or feature importance ranking techniques provided by machine learning models. Furthermore, ensuring the generalizability of the models across different social media platforms and user demographics presents another challenge. Addressing this requires collecting diverse datasets from multiple platforms and user groups, as well as incorporating transfer learning techniques to adapt models trained on one platform to others. Overall, overcoming these challenges necessitates a combination of data preprocessing techniques, advanced feature engineering methods, and model optimization strategies to develop robust and reliable fake profile detection systems.

VII. CONCLUSION

The implementation of a sophisticated hybrid model, integrating both deep learning and machine learning components, has shown promising results in detecting fake social media profiles. Key findings revealed that the K-Nearest Neighbors (KNN) classifier outperformed other models, achieving an accuracy of 84.12% after hyperparameter tuning. The proposed fusion model combining KNN with Multi-Layer Perceptron (MLP) and Long Short-Term Memory (LSTM) networks showed promise in enhancing detection accuracy and robustness. However, challenges such as the need for larger datasets and ethical considerations were acknowledged. Future research could explore techniques to address these challenges and further improve detection accuracy, potentially incorporating emerging technologies like natural language processing (NLP) for text analysis and graph neural networks for social network analysis. Additionally, investigating the impact of evolving tactics employed by malicious actors on detection systems could provide valuable insights for enhancing security and trustworthiness in online communities.

REFERENCES

1. Farooqui, Faisal and Usman Khan, Muhammed,” Automatic Detection of Fake Profiles in Online Social Network Using Soft Computing”, (July 18, 2022). International Journal of Engineering and Management Research, Volume-13, Issue-3 (June 2023), https://ssrn.com/abstract=4513674
2. G. Sansonetti, F. Gasparetti, G. D’aniello and A. Micarelli, "Unreliable Users Detection in Social Media: Deep Learning Techniques for Automatic Detection," in IEEE Access, vol. 8, pp. 213154-213167, 2020, doi: 10.1109/ACCESS.2020.3040604.
3. Social Networks Fake Profiles Detection Using Machine Learning Algorithms Innovations in Smart Cities Applications Edition 3, 2020ISBN : 978-3-030-37628-4 Yasyn Elyusufi, Zakaria Elyusufi, M’hamed Ait Kbir.
4. Ahmad, Shamim & Tripathi, Dr. (2023). A Review Article on Detection of Fake Profile on Social-Media. International Journal of Innovative Research in Computer Science and Technology. 11. 44-49. 10.55524/ijircst.2023.11.2.9.
5. F. Masood et al., "Spammer Detection and Fake User Identification on Social Networks," in IEEE Access, vol. 7, pp. 68140-68152, 2019, doi: 10.1109/ACCESS.2019.2918196.
6. L. Jiang and X. Zhang, "BCOSN: A Blockchain Based Decentralized Online Social Network," in IEEE Transactions on Computational Social Systems, vol. 6, no. 6, pp. 1454-1466, Dec. 2019, doi: 10.1109/TCSS.2019.2941650.
7. F. Yang, Y. Wang, C. Fu, C. Hu, and A. Alrawais, "An Efficient Blockchain-Based Bidirectional Friends Matching Scheme in Social Networks," in IEEE Access, vol. 8, pp. 150902 - 150913,2020,doi:10.1109/ACCESS.2020.3016986.
8. J. Kim and A. Yun, " Secure Fully Homomorphic Authenticated Encryption, " in IEEE Access, vol.9, pp.107279-107297, 2021, doi: 10.1109/ACCESS.2021.310084.
9. S.-C. Cha, T.-Y. Hsu, Y. Xiang, and K.-H. Yeh, “Privacy enhancing technologies in the Internet of Things: Perspectives andchallenges,” IEEE Internet Things J., vol. 6, no. 2, pp.2159–2187, April 2019.
10. B. Jeon, S. M. Ferdous, M. R. Rahman, and A. Walid, “Privacy Preserving decentralised aggregation for federated learning,” 2020, arXiv:2012.071836.
11. M. U. Hassan, M. H. Rehmani, and J. Chen, “Differential privacy in blockchain technology: A futuristic approach,” J.Parallel Distrib. Compute. vol. 145, pp. 50–74, Nov. 20197.

12. K. Yousaf and T. Nawaz, "A Deep Learning-Based Approach for Inappropriate Content Detection and Classification of YouTube Videos," in IEEE Access, vol. 10, pp. 16283-16298, 2022, doi: 10.1109/ACCESS.2022.31475198.
13. M. -S. Kang, R. -H. Park and H. -M. Park, "Efficient Spatio-Temporal Modelling Methods for Real-Time Violence Recognition," in IEEE Access, vol.9, pp.76270 - 76285, 2021, doi: 10.1109/ ACCESS.2021.3083273.
14. P. Sharmila, K. S. M. Anbananthen, D. Chelliah, S. Parthasarathy and S. Kannan, "PDHS: Pattern-Based Deep Hate Speech Detection with Improved Tweet Representation," in IEEE Access, vol. 10, pp. 105366-105376, 2022, doi: 10.1109/ACCESS.2022.3210177
15. J. Sun, X. Yao, S. Wang, and Y. Wu, "Blockchain-Based Secure Storage and Access Scheme For Electronic Medical Records in IPFS," in IEEE Access, vol. 8, pp. 59389-59401, 2020, doi: 10.1109/ACCESS.2020.298296.