e-ISSN:2582-7219

INTERNATIONAL JOURNAL OF MULTIDISCIPLINARY RESEARCH IN SCIENCE, ENGINEERING AND TECHNOLOGY

Volume 7, Issue 12, December 2024

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

Impact Factor: 7.521

© 2024 IJMRSET | Volume 7, Issue 12, December 2024| DOI: 10.15680/IJMRSET.2024.0712130

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 7.521| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

Identifying Fake Accounts on Facebook using Machine Learning

Sushil Mahato, Aryan Dipak Raut, Sushil Sapkota, Rakesh Kumar Nayak

Department of Computer Science & Engineering, Visvesvaraya Technological University, Bangalore, India

ABSTRACT: In modern digital era, online social networks (OSNs) have become necessary for communication, networking, and information sharing among the people throughout world.

This hasmade their general use one of the favored targets of cyber criminals to cause malicious activities. These activities include the generation and spread of fake accounts on Facebook. The current research work covers the identification of fake accounts in the most widely used OSN, Facebook. This three-phase approach covers data collection, identification of features, and machine learning-based classification.

The data collection phase involved gathering information on both genuine and fake Facebook accounts. Feature identification utilized user feed data to analyze activity patterns and define five critical characteristics that differentiate fake accounts from real ones. These characteristics were then applied in machine learning classifiers, including K-Nearest Neighbours (KNN), Support Vector Machines (SVM), and Neural Networks (NN).

The results revealed that the KNN classifier achieved the highest precision, with an accuracy rate of 82% in total. Key findings indicate that "likes" and "remarks" significantly contribute to the detection process. While the precision is not flawless, the study underscores the ability of fake accounts to mimic real users. These findings highlight the importance of machine learning in combating cyber threats and enhancing the cyber security of OSNs.

KEYWORDS: Fake Account Detection, Facebook Account Identification, Data Mining for Fake Accounts, KNN for Fake Account Detection, Account Cloning Detection

I. INTRODUCTION

Online social networks (OSN) have nowadays become an essential part of daily life, serving as common platform for communication and information sharing. Different websites like Facebook, Twitter, Instagram, and many others are being used to update one's status, pictures, and videos, among others, to share information, enabling seamless interaction even over distant locations. According to the 2023 global data, 59.4% of the world population uses online social networks, but the number is growing rapidly.

The essence of OSNs lies in their association with user identity. As noted by researchers, identity is a distinct characteristic linked to an individual, often represented by names or unique identifiers such as passports. A passport, for example, contains crucial personal information, including a person's name, date of birth, address, nationality, national ID, and even biometric data like fingerprints. While individuals may have multiple identities in specific contexts, each must be unique to its owner.

However, the misuse of false identities in OSNs has become a growing concern. Such identities are often created to impersonate others with malicious intent, such as harvesting personal information for targeted cyberattacks, spreading propaganda, or conducting fraudulent activities like phishing, spamming, and scamming. This misuse directly contradicts the fundamental purpose of social networks, which is to foster genuine connections.

Modern OSNs, including Instagram and Twitter, provide robust Application Programming Interfaces (APIs) that allow real-time and accurate data acquisition. Facebook, one of the most widely used platforms globally, also offers APIs to access user profile data, covering static and dynamic components. Static data includes user-defined details like age,

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 7.521| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

relationship status, and preferences, while dynamic data reflects user activity, social interactions, and demographic information. Machine learning techniques such as K-Nearest Neighbours (KNN), Support Vector Machines (SVM), and Neural Networks (NN) are increasingly employed to analyze such data, enabling the detection of patterns indicative of fake accounts.

This study focuses on identifying fake accounts on Facebook, particularly within Southeast Asia, where user demographics often display similarities in terms of language, culture, and usage patterns. Given Facebook's popularity in Malaysia and the region as a whole, this research leverages machine learning techniques to examine user behavior and profile data, aiming to enhance the detection of fraudulent accounts and address this pressing cybersecurity challenge.

II. LITERATURE REVIEW

To access the services provided by online social networks (OSNs), users are typically required to create a Facebook profile containing essential personal data, which may include name, gender, location, email address, and other essential private information. The free to use of these platforms, while facilitating communication and connection, also make them vulnerable by malicious cyber criminals. These adversaries create fake profiles to carry out various unlawful, malicious, or deceptive activities such as spamming, marketing, stalking, defamation, and other harmful actions. The motivations for creating false profiles often depend on the specific social network targeted. For example, on platforms like Facebook and Twitter, attackers may aim to steal private information, promote a specific brand or individual, or engage in character assassination. Similarly, on professional networks such as LinkedIn and ResearchGate, adversaries may generate fake profiles to monitor user behaviour or manipulate experts in specialized fields. On dating platforms, attackers, often referred to as "Catfishers," may create fake profiles to exploit vulnerable individuals seeking romantic relationships, often leading to financial exploitation or emotional manipulation.

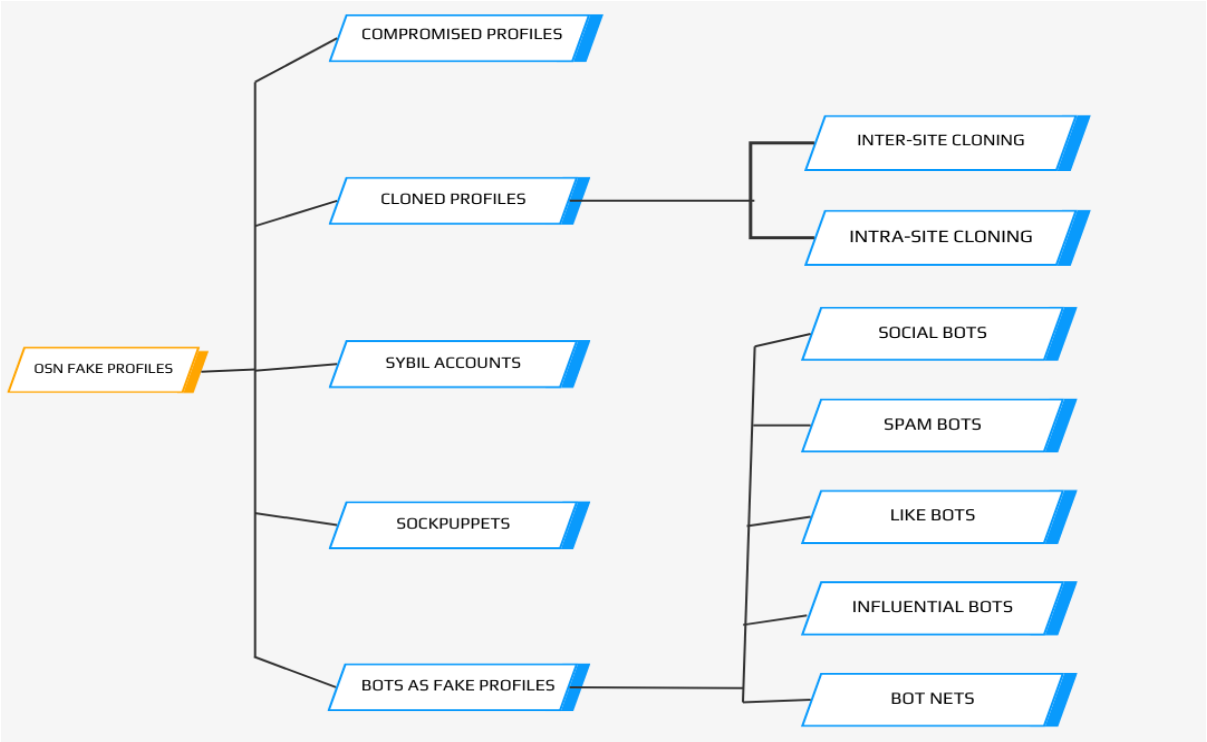

Figure 1: Type of false online social network profile

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 7.521| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

According to previous research by, fake profiles on OSNs can be categorized into five distinct types: compromised profiles, cloned profiles, Sybil accounts, sock puppets, and fake bot profiles. Cloned profiles, for example, can be further divided into inter-site cloning (where a fake profile is created on multiple platforms) and intra-site cloning (where the profile is duplicated within the same platform). Similarly, fake bot profiles encompass various types of automated bots, such as social bots, spam bots, like bots, influential bots, and botnets. These categories represent the different tactics employed by adversaries to achieve malicious objectives across various online social networks, as illustrated in Figure 1.

To detect fake accounts, scholars have identified several key attributes and features that distinguish these profiles from genuine ones. Accurate and effective detection of fake accounts requires the careful selection of relevant features, which can either be manually observed on OSNs or derived from existing literature. However, it is important to note that some characteristics may lose their effectiveness over time as attackers continuously evolve their strategies to bypass detection systems. Numerous studies have explored different characteristics of social network profiles in order to develop and refine detection models for fake accounts. Based on the nature of these attributes, this study categorizes them into five main groups:

1. **Network-based Attributes**

These attributes focus on the relationship structure between users and their connections. By analyzing the degree of connection (e.g., first-degree connections such as friends and second-degree connections like friends of friends), it is possible to identify unexpected behaviour that may indicate the presence of a fake profile.

2. **Content-based Attributes**

Content-based analysis involves checking the types and quality of content shared by users on the online social platforms. Irregular or suspicious content patterns, such as excessive sharing of promotional material or spam-like posts, can serve as indicators of fake social accounts.

3. **Temporal Features**

Temporal features relate to the timing and frequency of activities performed by a user. A profile exhibiting irregular or unnatural patterns of activity, such as frequent login times or activity bursts at odd hours, may indicate the presence of a fake account.

4. **Profile-based Features**

Profile-based features focus on the behaviour of the user's profile, such as the number of people followed, frequency of posts, and other profile interactions. Unclear in these features, such as a large number of followers with little to no engagement, can signal a fake Facebook account.

5. **Action-based Features**

These features examine the types of actions taken by the user on the platform, including tagging behaviour, location sharing, and friend tagging. Unusual patterns in these actions, such as over-tagging or automated posting behaviours, can also be indicative of a fake profile.

By analyzing these various attributes and features, we are able to detect malicious activities that deviate from typical user behaviour on Facebook. These deviations play a crucial role in the decision-making process during the classification of fake accounts, helping to create more accurate and efficient detection models. As adversaries continue to refine their tactics, these attributes must be continually adapted and updated to maintain the effectiveness of fake account detection systems.

III. METHODOLOGY

3.1 Data Collection Method

For this study, real-world Facebook datasets are essential; however, such datasets are not openly accessible to the public. While some social graph datasets exist that include profile-based feature data, they are anonymized and unavailable for use. As a result, this study relies on data obtained through the Facebook API, which is subject to authorization restrictions. This challenge is commonly noted by researchers working with Facebook data, as highlighted by. Furthermore, Facebook's frequent updates to its security and privacy policies make it difficult to access data without explicit permission from the platform itself. The data collection process, as depicted in Figure 2, involves both API-based and bot-crawler approaches. These techniques are time-consuming and highly sensitive to the privacy settings and security protocols of users. To address the challenge of collecting real Facebook data, this study generates

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 7.521| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

synthetic data samples based on the network structure and characteristics of available datasets. Synthetic data can be created using different tools based on the statistical parameters of current social networks, such as degree distribution, clustering coefficient, and centrality measures. Several online data generators, such as GEDIS Studio, Databene Benerator, and Mockaroo, offer the ability to produce such artificial datasets. After careful evaluation, Mockaroo was selected as the data generation tool due to the realism of its output, which closely resembles real user data. A total of 800 synthetic Facebook user data samples were generated, consisting of both real and fake accounts. The details of the collected Facebook user data are summarized in Table I.

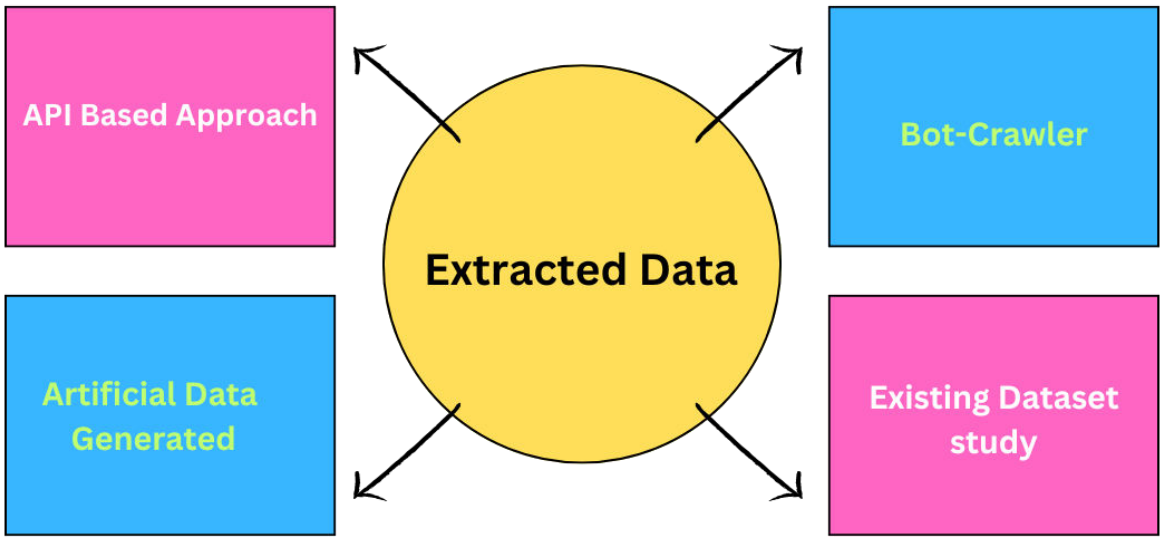

Figure 2: Data Collection Technique

Serial	Description	Value
1	Total Users	800
2	Real Users	615
3	Fake Users	185
4	Assumed Real Users	560
5	Assumed Fake Users	240

Table I: Facebook User Data Collection

3.2 Feature Identification

Following the data collection, the next step involves identifying and defining a set of features that can help distinguish real users from fake users. A set of 17 potential characteristics was initially selected from various literature sources, including. After further refinement based on studies by, the most significant features for detecting fake accounts were selected. These features are presented in **Table II** along with their descriptions and the rationale behind their inclusion.

© 2024 IJMRSET | Volume 7, Issue 12, December 2024| DOI: 10.15680/IJMRSET.2024.0712130

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 7.521| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

Serial	Feature Name	Feature Description	Justification	Measuring Methodology
1	Average Post Likes Received	Average number of likes a user receives on their own posts	Fake accounts tend to share spam-like content, resulting in fewer likes.	Data can be extracted from posts in the user's feed.
2	Average Post Likes	Average number of posts liked by the user per day	Fake accounts exhibit higher activity, unlike real users who show more balanced engagement.	Can be measured by counting the posts the user likes in their feed.
3	Average Post Comments Received	Average number of comments on a user's own posts	Fake accounts tend to post spam with minimal engagement, leading to fewer comments.	Comment data can be gathered from user posts.
4	Average Post Comments	Average number of comments made by the user per day on posts	Fake accounts may exhibit abnormally high commenting behavior on spam content.	Count the number of comments made by the user on posts per day.
5	Average Friends	Average number of friends connected to the user	Fake accounts may have an unusually high number of friends compared to real users.	Data can be gathered by examining the user's friend list.

Table II: Feature Set with Description and Justification

3.3 Learning Classifiers

In the final phase of this methodology, supervised machine learning algorithms are employed to classify Facebook accounts as either real or fake. Supervised learning algorithms rely on labelled datasets to build predictive models that can make classifications based on input features. In this context, the two classes of users—real and fake—are the primary focus. The assumption underlying the use of supervised learning classifiers is that the feature values of genuine users and fake accounts will differ significantly, particularly when fake accounts engage in anomalous behaviours.

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 7.521| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

3.3.1 K-Nearest Neighbour (KNN)

The K-Nearest Neighbour (KNN) algorithm will be in use for classification based on the proximity in feature space to training examples. This is one of the simplest machine learning algorithms, in which the training data consists of vectors with features and labels. In this algorithm, a query point is assigned a label which is the majority label of the k nearest neighbours in the feature space. This method is straightforward yet effective in highlighting patterns of user behaviour that demarcate real users from fake accounts.

3.3.2 Support Vector Machine (SVM)

Support Vector Machine (SVM) is another classification technique that finds a hyperplane in feature space to separate instances of different classes. SVM operates by mapping data into higher-dimensional space and identifying the optimal hyperplane that best divides the data points into two categories. It is particularly effective in handling both continuous and categorical variables and is widely used in classification tasks. SVM's robustness to high-dimensional data makes it well-suited for distinguishing between real and fake Facebook accounts based on complex feature sets.

3.3.3 Neural Network (NN)

Neural networks (NN) are computational models inspired by the human brain, consisting of interconnected nodes or neurons that process and transmit information. NN algorithms learn patterns by adjusting the weights of the connections between nodes so as to minimize the error in the predictions. These models contain input, hidden, and output layers; the hidden layers internally perform computations to refine the network's output. NN is especially good when the recognition of complicated patterns is required; it is a good fit for fake account detection that relies on complex behaviour patterns.

IV. EVALUATION OF DETECTION TECHNIQUES

This section evaluates the effectiveness of various techniques used to identify fake accounts on Facebook. The classifiers discussed earlier were applied to a blended dataset, which included previously identified real accounts from both the first and second stages of users within the social neighbourhood, alongside fake accounts. Additionally, the dataset comprised user accounts of friends from the social neighbourhood, with assumed real accounts for active users and assumed fake accounts for inactive users. The study aimed to assess the performance of different machine learning classification models—KNN, SVM, and NN—using the Orange tool to predict unknown user accounts.

Dataset Preparation and Clustering

The initial step involved dataset cleaning to prepare it for classification. In this phase, clustering methods—Linear Projection and Circular Placement—were applied to the dataset using the clustering feature in the Orange tool. The combination of features listed in Table II was used to categorize the data into four distinct clusters:

- **G1**: Fake account users
- **G2**: Assumed fake account users
- **G3**: Inactive users
- **G4**: Real users

Classifier Training and Evaluation

Following the clustering phase, the dataset was subjected to the classifier learning phase, where three distinct techniques—KNN, SVM, and NN—were used to assess the ability of each model to detect fake accounts. According to prior research, these three classifiers (KNN, SVM, and NN) have demonstrated effectiveness in fake account detection on social media platforms like Facebook. This comparison allowed for an evaluation of which learning classifier yielded the best results.

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 7.521| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

Table III presents the evaluation metrics for each classifier, showing the performance of each model based on key criteria:

Serial	Model	Area Under ROC Curve	Classifier Accuracy	Balance of F-Score	Precision
1	KNN	0.967	0.829	0.781	0.760
2	SVM	0.794	0.729	0.685	0.665
3	NN	0.958	0.800	0.777	0.772

Interpretation of Results

Classifier Accuracy (CA) represents the proportion of correct predictions made by each model. A higher CA value indicates a stronger predictive model. As shown in **Table III**, the KNN model achieved the highest CA (0.829), outperforming both the SVM and NN models. The area under the ROC curve (AUC) was highest for KNN (0.967), signifying its superior ability to distinguish between real and fake accounts. The balance of the F-Score and precision values also supported the KNN model's effectiveness.

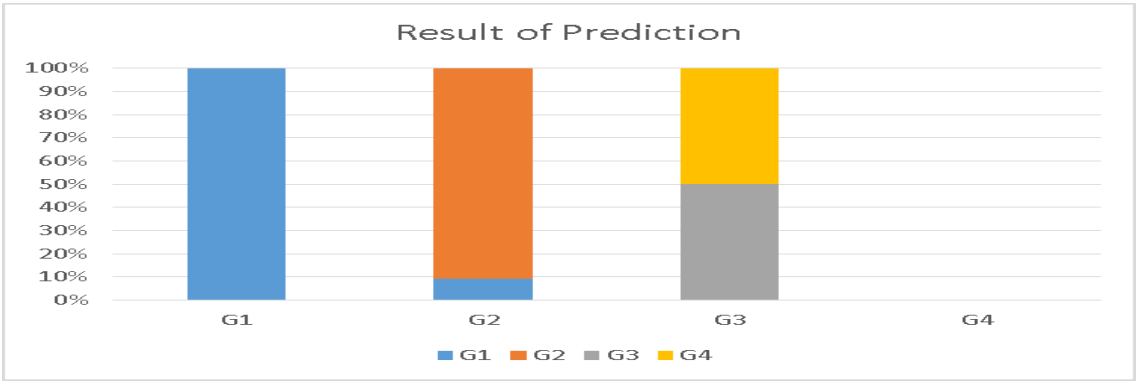

Figure 3: Prediction result for KNN

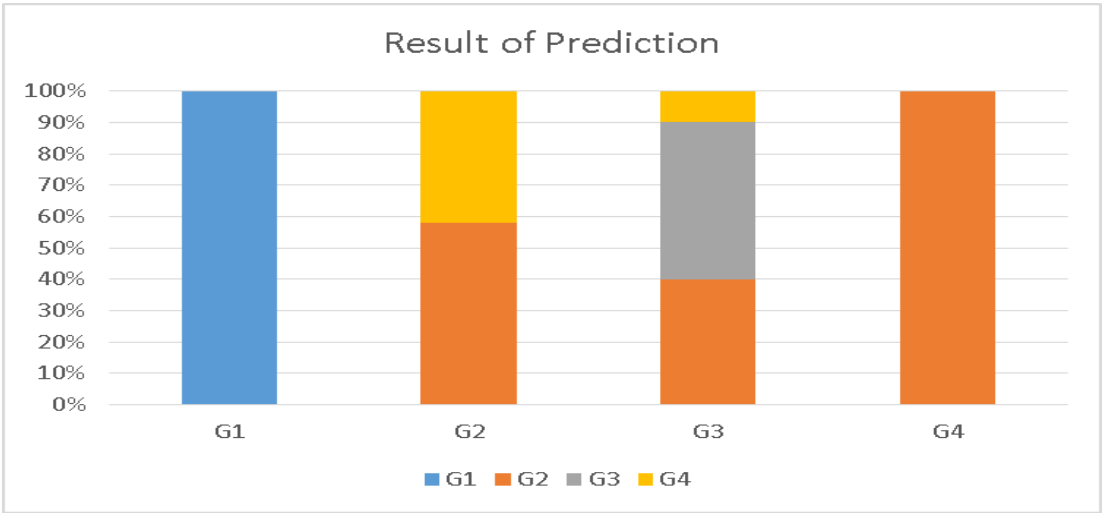

Figure 4: Prediction Result for SVM

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 7.521| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

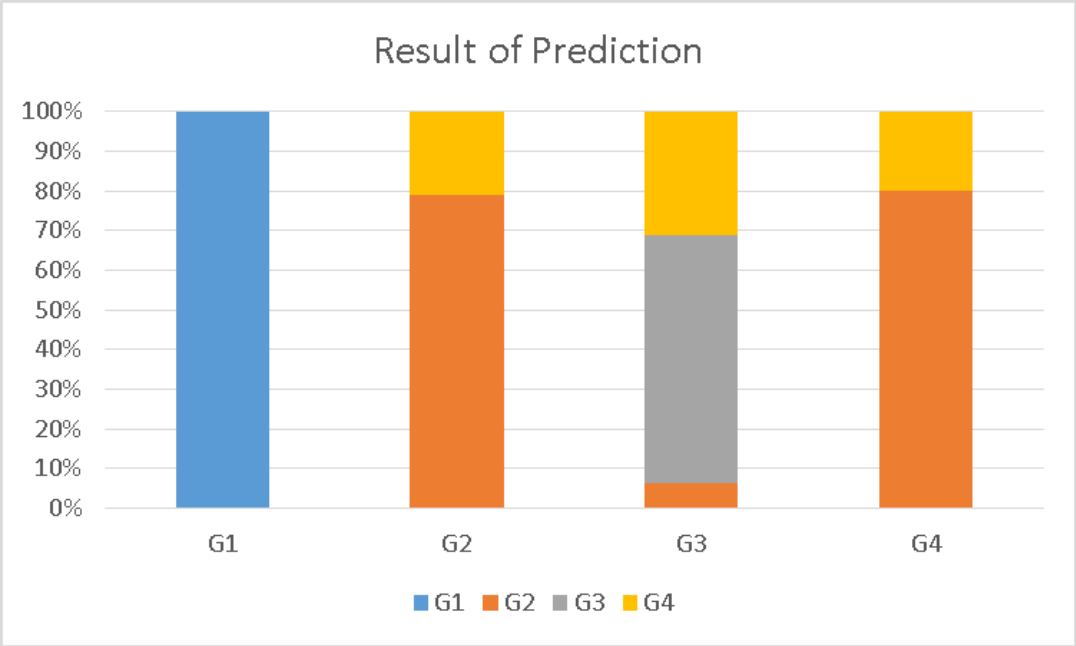

Figure 5: Prediction Result for NN

Figures 3, 4, and 5 illustrate the prediction results for each classifier.

- **Figure 3**: KNN model's prediction results show that almost all fake accounts in the G1 group were correctly identified, with detection accuracy up to 70% for the G2, G3, and G4 groups.
- **Figure 4**: The SVM model's prediction performance mirrors that of KNN, with up to 70% detection accuracy across all groups.
- **Figure 5**: The Neural Network (NN) model also achieved 70% detection accuracy for all groups, supporting its reliability as a detection tool.

V. CONCLUSION

The detection of fake accounts on platforms like Facebook has become increasingly challenging as these accounts continuously adapt to evade identification. This study focused on the identification of fake accounts specifically in Southeast Asia, leveraging artificial datasets due to the complexities involved in collecting real user data, particularly because of Facebook's fine-grained privacy settings. By applying widely recognized machine learning classification methods, the study aimed to identify the most effective classifiers for detecting fraudulent accounts.

The findings underscore the importance of refining detection techniques and exploring innovative approaches. Future research should explore hybrid models that combine multiple techniques to enhance accuracy. Additionally, the inclusion of other relevant data points—such as account IDs, location information, and device usage patterns—could provide more robust solutions for fake account detection. Such advancements would greatly contribute to improving the security and integrity of online social platforms in the future.

REFERENCES

1.A. Romanov, A. Semenov, O. Mazhelis, and J. Veijalainen, "Detection of Fake Profiles in Social Media Literature Review," no. Webist, pp. 363–369, 2017.
2.A. Kumbhar, M. Wable, S. Nigade, K. Darekar, and B. E. Student, "A survey on: Malicious Application and Fake user Detection in Facebook using Data Mining," Int. J. Eng. Sci. Comput., vol. 7, no. 12, p. 15768, 2017.

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 7.521| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

3.A. Guess, J. Nagler, and J. Tucker, "Less than you think: Prevalence and predictors of fake news dissemination on Facebook," Asian-Australasian J. Anim. Sci., vol. 32, no. 2, pp. 1–9, 2019.
4.P. S. Rao, J. Gyani, and G. Narsimha, "Fake Profiles Identification in Online Social Networks Using Machine Learning and NLP," Int. J. Appl. Eng. Res. ISSN, vol. 13, no. 6, pp. 973–4562, 2018.
5.M. B. Albayati and A. M. Altamimi, "An empirical study for detecting fake facebook profiles using supervised mining techniques," Inform., vol. 43, no. 1, pp. 77–86, 2019.
6.M. Fire et al., "A sneak into the Devil's Colony - Fake Profiles in Online Social Networks," J. Supercomput., vol. 5, no. 1, pp. 26–39, 2018.
7.A. M. Ali, H. Alvari, A. Hajibagheri, K. Lakkaraju, and G. Sukthankar, "Synthetic Generators for Cloning Social Network Data," BioMedCom, pp. 1–9, 2014.
8."Facebook Data Policy," 2018. [Online]. Available: https://www.digitaltrends.com/social-media/terms-conditions-facebooks-data-use-policy-explained/accessed. [Accessed: 16-Aug-2019].
9."Software Testing Help: Top 10 Best data Generatools in 2019." [Online]. Available: https://www.softwaretestinghelp.com/test-data-generation-tools. [Accessed: 14-Aug-2019].
10."Generated Data: Generated Data about." [Online]. Available: https://www.generatedata.com/#t2. [Accessed: 15-Aug-2019].
11."No Title9. Mockaroo Realistic Data Generator:" [Online]. Available: https://mockaroo.com/. [Accessed: 15-Aug-2019].
12.A. Gupta and R. Kaushal, "Towards detecting fake user accounts in facebook," ISEA Asia Secur. Priv. Conf. 2017, ISEASP 2017, vol. 1, pp. 1–6, 2017.
13.R. Feizy, "An evaluation of identity in online social networking: distinguishing fact from fiction," 2010.
14.S. Gheewala and R. Patel, "Machine Learning Based Twitter Spam Account Detection: A Review," in Proceedings of the 2nd International Conference on Computing Methodologies and Communication, ICCMC 2018, 2018, no. Iccmc, pp. 79–84.
15.A. M. Likhon, A. S. M. M. Rahman, and M. H. Choudhury, "Detection of fake identities on twitter using supervised machine learning," Brac University, 2019.
16.J. Kim, B.-S. Kim, and S. Savarese, "Comparing Image Classification Methods: K-Nearest-Neighbor and Support Vector-Machines," Appl. Math. Electr. Comput. Eng., pp. 133–138, 2012.
17.S. Kudugunta and E. Ferrara, "Deep neural networks for bot detection," Inf. Sci. (Ny)., vol. 467, pp. 312–322, 2018.
18.R. Raturi, "Machine Learning Implementation for Identifying Fake Accounts in Social Network," vol. 118, no. 20, pp. 4785–4797, 2018.

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

INTERNATIONAL JOURNAL OF MULTIDISCIPLINARY RESEARCH

IN SCIENCE, ENGINEERING AND TECHNOLOGY

| Mobile No: +91-6381907438 | Whatsapp: +91-6381907438 | ijmrset@gmail.com |

www.ijmrset.com