Journal of Data Acquisition and Processing
ISSN: 1004-9037
https://sjcjycl.cn/
DOI: 10.5281/zenodo.7766430

FAKE ACCOUNT IDENTIFICATION USING MACHINE LEARNING APPROACHES INTEGRATED WITH ADAPTIVE PARTICLE SWARM OPTIMIZATION

Dhruvi Patel[1,a],Dr. Chintan Thacker [1,b], Dr. Kamal Sutaria [1,c], Hetal Bhaidasna[1,d]

[a] Student of Computer Science Department, Parul Institute Of Engineering and Technology, Limda, Vadodara, Gujarat, India
[b] Assistant Professor Of Computer Science Department, Parul Institute Of Engineering and Technology, Limda, Vadodara, Gujarat, India
[c] Assistant Professor Of Computer Science Department, Parul Institute Of Engineering and Technology, Limda, Vadodara, Gujarat, India
[d] Head Of Department Computer Engineering-Diploma Studies, Parul Institute Of Engineering and Technology, Limda, Vadodara, Gujarat, India
[a)]dhruveepatel51@gmail.com
[b)] chintan.thacker19435@paruluniversity.ac.in
[c)]kamal.sutaria24554@paruluniversity.ac.in
[d)]hetal.bhaidasna@paruluniversity.ac.in

***Abstract*—** It is common practice for humans, bots, and other automated systems to create new user accounts using stolen or otherwise fraudulent personal data. They are used in deceptive practices like as phishing and identity theft as well as in the propagation of malicious rumors. A single malicious actor may create hundreds or even thousands of fake accounts in order to spread their malicious activities to as many real users as feasible. Users may learn a great lot from social networking platforms. Hackers have an open invitation to exploit this trove of social media data. These cybercriminals create false personas and spread pointless content. There is a critical step in navigating social media networks that involves identifying fake profiles. In this research, we offer a machine learning method for spotting Instagram phoney profiles. We used the attribute-selection technique, adaptive particle swarm optimization, and feature-elimination recursion in this strategy. The findings show that the proposed adaptive particle swarm optimization approach outperforms RFE in terms of accuracy, recall, and F measure.

***Keywords*—** Fake account, Machine Learning, Feature Selection, Adaptive Particle Swarm Optimization, Recursive Feature Elimination.

INTRODUCTION

With millions of users and billions of minutes spent on such sites, social networking has developed into a well-known pastime on the internet today. Online social network services range from social interaction-based platforms resembling Facebook or Myspace to knowledge dissemination-focused platforms resembling twitter or Google Buzz to social interaction features added to current systems like Flicker. On the other side, improving security issues and maintaining OSN privacy continue to be a top priority and viewed mission. The use of social

media platforms is widespread for sending and receiving data. Everyone utilizes social media, whether it's to post pricey images, follow celebrities, or communicate with close and far friends. It is an excellent setting for socializing and exchanging information. [1].They utilise them to interact with the content and information provided by other users on the network. Social media has enabled people from all walks of life to connect with one another and exchange information and creative works. It also provides a platform for individuals to showcase their skills and make connections with others all around the globe who have similar passions [2]. Social media comes in many forms, including Instagram, LinkedIn, Twitter, Facebook, snapchat, and others. Social networking has become a daily requirement for the majority of people as technology advances.

A fake account is basically an account on any social media platform where the details displayed are actually dishonest oreven fraudulent. Misrepresentation on fake accounts, using false details, deceives the general public to spread inaccurate information or collect financial or personal information [8]. People set up accounts on various social networking sites to exchange social media content. In order to propagate false information without disclosing their identities, users frequently create accounts with false or anonymous information. Users frequently alter their accounts or create accounts in someone else's name (identity theft). There are certain specific financial advantages to creating false accounts. These phone identities are maintained by bots or automated programs, which aid in the deeper and faster dissemination of fake news online. On the network, fake accounts frequently connect and follow the posts of influencers. Through policies against impersonation, even social media like twitter, Facebook, and WhatsApp that are used for online social networking erase or freeze these phone accounts. Within hours of the explosion during the Boston Marathon, approximately 32,000 new accounts were created, of which twenty percent had their profile discontinue by twitter (Gupta et al. 2013). The majority of malicious profiles are created with the intention of spamming, phishing, and getting more followers. The fake accounts have all the tools necessary to commit online crimes. Identity theft and data breaches are substantial risks posed by fake accounts. All user information is transferred to faraway servers when users view the URLs sent by this fake profile, where it may be utilized against them. False profiles that claim to be from companies or individuals can harm their reputation and get them less likes and follows [9].

LITERATURE STUDY

In [5], authors recommend a detection technique as 3PS (Publicly Private Protected system) for detecting a fake account over an online social network by considering the behavior patterns of user account activity. It emphasizes the detection of fake or malicious users who users own various online social network accounts, send friend requests, and share many posts as malicious. It considers the posts, follows, status updates, followers, and posts. A fake or malicious user can be detected by distinguish between the threshold value of an attribute related to the user's personal profile along with network similarity exploration.

In [6] a system to identify a fake account on twitter by selecting the user profile characteristic feature and logistic regression with PSO, naïve bayes and KNN algorithms used for the classification task. They used information gain, correlation-based feature selection (CFS), minimum relevance maximum redundancy (MRMR) for feature selection. The dataset, which

consisted of 6973 profile data and was split into training and testing sections, was gathered manually and in other ways utilizing the Twitter REST API.

In [7], used a framework that performs a dual analysis: message reliability prediction and social network user profile reliability prediction. It includes both offline analysis with deep learning algorithms included and online analysis with actual users to classify a Twitter profile as trustworthy or not. The United Nations should implement a method suggested by S. Uppada et al. [8]. By rating users and articles, as well as fake profile identification is utilized as a statistic to track user engagement trends. In order to create characteristics related to fake news images, forensic techniques and picture polarity analysis are merged. The SENAD approach and the CredNN model each had accuracy levels of 76.3% and 93.5% for fake news, respectively.

In [9] "researcher recommends a new bot identification methodology using deep neural networks and active learning. Modules for feature extraction, active learning, data gathering and labeling, and detection are included. In addition, this framework offered a most recent RGA deep neural network model for identification that makes use of ResNet, BiGRU, and attention mechanisms. The testing findings demonstrated that the suggested DA Bot framework is superior to existing detection methods in terms of effectiveness for identifying social" bots.

In [10] uses a framework which is based on chrome extension that identifies a fake twitter account. They also contrasted a number of tasks utilizing machine learning strategies to confirm user information gathered through manual and web crawler. The data set was obtained from the twitter profile using a web crawler and the twitter API. The profile data gathered includes the user's name, ID, number of status updates, friends list, number of favorites, and number of URLs cited in tweets. Then the data is grouped into test and training data in an 80:20 ratio. And passed the WEKA machine learning platform. The trust score from the features is calculated and then goes to the Chrome extension to identify the malicious profile. Through user distinctive analysis and the usage of a trust score, Chrome extensions produce a suspiciousness score for each user. They employ random forest and bagging to identify fraudulent profiles and generalize the performance of the Chrome plugin.

In [11] A semi-supervised clustering strategy that can account for partial background information, in conjunction with the deep walk method on rater graphs, provides a top-down framework for detecting sham groupings of potential reviewers. Additionally, using temporal affinity, semantic traits, and sentiment analysis, this strategy can be expanded to find groups of opinion spammers on social media. Detection of potential spammer groups based solely on the underlying graph's topological structure. The reviewer ID, represented by feature vectors, is divided into various groups of spammer candidates using a modified version of the semi-supervised clustering process known as Pair wise after first obtaining a representation for each node in G using the Deep Walk method. Confined K-Means. To identify groups of fraud reviewers from reviewer graphs, the framework was verified on a partial core data set consisting of 2207 fraudulent reviewer IDs belonging to 23 different reviewers.

In [12] suggest a technique using dynamic knowledge graphs to find fake feedback. Using a newly developed neural network model termed conditional two-way long-term short-term memory, the first four types of entities are extracted by embedding phrase vectors/double words based on the characteristics of online product reviews. Next, time series-related properties are incorporated into the knowledge graph building process to generate dynamic

graph networks. They first extracted many different sorts of entities using the ST-BLSTM algorithm, after which they provided the most recent MI-based criteria for evaluating the connections between entities. They subsequently produced a dynamic graph network by mixing the time sequence factors. The second showed the value of evaluating the truthfulness and commodity quality of the rather by calculating trustworthiness, honesty, great commodity degree, and reliability ratings utilizing novel methodologies. The technique demonstrated how review portrait data reflects the reasons of bogus reviews and provides crucial information. The technique demonstrated how review portrait data reflects the root reasons of false reviews and offers crucial hints regarding various reviewer types.

In [20], propose long-short memory neural network and an AdaBoost model are combined in the geolocation-based profile recognition prototype to analyze user account and geolocation data. Its two sub-models are the geolocation detection model and the profile detection framework. Geolocation DM receives geolocation features and applies LSTM to analyze geolocation sequences to create user identity prediction scores. Account-DM receives account feature input and analyses account feature using AdaBoost. The linear classifier SVM generates the final evaluation of user IDs using the profile-DM and geolocation-DM prediction scores as input. This technique may reliably and effectively identify fraudulent reviews, according to the study, which used a huge dataset from Yelp.

PROPOSED SYSTEM

On this paper we presented a different machine learning approaches and attribute selection techniques used to observe the false profiles in online social media. Moreover, we are adding the Adaptive Particle Swarm optimization and recursive feature elimination techniques to increase the detection accuracy rate of the fake profiles.

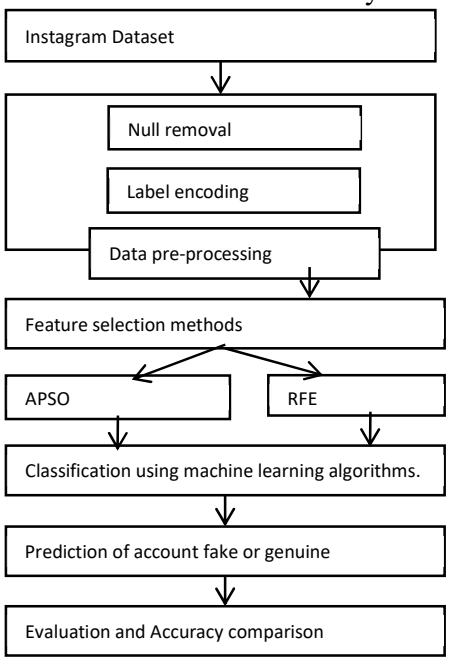

Fig. 1. Proposed System Flow-chart

The following is a summary of the steps for identifying fake profiles:

1.Collect the Instagram dataset from Kaggle.
2.Data pre-processing includes removing null data, label encoding.
3.The data set should be divided into test and training data.
4.Applying feature elimination techniques APSO and RFE.
5.Predict the test result of fake and genuine profiles.
6.Compute accuracy.
7.Compare the performance of classification models.

A. Dataset

Instagram accounts authenticity has been determined using data from a Kaggle competition. There are ten different types of information in this dataset, including followers, following, post count, username length, profile image, description length, private, public, and external URL. There are 13 fields to explore in the Kaggle dataset. These characteristics may be used as clues to determine whether a profile is real or a fake. In order to train the machine learning models, all of the characteristics had to be investigated and assessed.

TABLE I. ATTRIBUTES FOR INSTAGRAM PROFILE

Attributes	Attributes Description
Profile pic	Profile picture is present or not
nums /length username	Ratio of number of numerical chars in username to its length
full name words	full name in word tokens
name==username	Username and full name are same
nums/length full name	Ratio of number of numerical characters in full name to its length
description length	Length of the description in account
external URL	Any link or URL is available or not
private	Profile is private or public
posts	Number of posts
follower	Number of Follower
following	Number of following

B. Pre-processing

Incomplete, inconsistent, and missing in particular behaviours or patterns, real-world untrained data are common. They may also be full of mistakes. This means that the machine learning algorithm may utilise the data for the model after they have been gathered and prepared.
Disposable Zero - One of the crucial actions in data wrangling is the elimination of null values.

Any machine learning method will suffer from decreased performance and accuracy due to these missing variables. Therefore, prior to using any machine learning technique, it is crucial to clean the dataset of any null values.

Popular techniques of encoding categorical variables include label encoding. Each label is assigned a unique number following the established system of alphabetical order.

C. Feature selection technique

The search efficiency of the given adaptive particle swarm optimization (APSO) is higher than that of the traditional particle swarm optimization. Additionally, it is capable of a quicker convergence rate and a global search throughout the whole search space. The APSO has two primary phases. First, a real-time evolutionary state estimate approach is carried out to identify one of the four designated evolutionary stages, namely exploration, exploitation, convergence, and leaping out in each generation by analysing the population distribution and particle fitness. It provides real-time adaptive tuning of search efficiency and convergence rate by adjusting algorithmic parameters such as inertia weight and acceleration coefficients.

This new approach creates an initial population that contains solutions distributed uniformly across all segments. The proposed algorithm's search capability is enhanced by segmenting the entire search space. The proposed model strengthens the swarms' ability to share information. Every swarm expresses interest in or gathers information from other swarms that are more fit than it is. The following steps are carried out for the algorithm:

1. Set up the PSO parameters, including the maximum number of iterations, population size, and initial particle positions and velocities.
2. Define an objective function that calculates the fitness of each particle based on its feature subset.
3. Evaluate the fitness of each particle using the objective function.
4. Identify the particle with the best fitness as the global best particle.
5. Begin the main PSO loop.
6. Update the velocity and position of each particle using the PSO update equations.

The velocity update of algorithm is as follows:

```
Vi = Vi*wi + 1 / rank(i)*rand()*(pbest[i] – Xi) +        AdaptivePSO();
    Xi = Xi + Vi;
        Where,
        AdaptivePSO(i)
        {
        Posx← 0.0
        For each individual k of the population
                if pFitness[k] is better than fitness[i]
posx←posx + 1 / rank(k) * rand()*(pbest[k]-Xi);
        if(pos>Vmax)return Vmax:
        else return posx;
}
```

Where, pfitness[k] represent the best local fitness, fitnessì] indicates the current fitness of swam.TheAdaptivePSO() module gives the direction of the swarm by sharing information with

all other individuals that have better fitness,Vmax have been set with a small value to prevent jump.

7. Evaluate the fitness of each particle based on its new position, and update its personal best and the global best particle if necessary.
8. Calculate the swarm diversity using a diversity measure.
9. If the diversity is below a certain threshold, increase the exploration probability and decrease the exploitation probability.
10. Update the PSO parameters based on the exploration and exploitation probabilities.
11. Continue the loop until the maximum number of iterations is reached or a stopping criterion is met.
12. Select the best feature subset found by the PSO algorithm based on the global best particle.

To pick the best features for a model, recursive feature elimination (RFE) first finds the optimal number of features and then iteratively eliminates the features with the lowest predictive value. Using the model's attributes to rank features, RFE then tries to get rid of any dependencies and collinearity in the model by iteratively dropping a few features at a time. For RFE to function, a minimum threshold of valid features must be met, however this threshold is not always known in advance. Cross-validation is used in conjunction with RFE to score several feature subsets and pick the highest scoring collection of features, allowing for the determination of the optimum amount of features.

D. Classification

Classification is a technique used in machine learning that allows classes to be learned and applied to a problem. In the study being suggested, an evaluation is made on whether or not the profiles in question are authentic. As part of the suggested effort, a value of 1 indicates a phoney profile, whereas a value of 0 indicates the opposite. Support Vector Machine, KNN, Random Forest, Logistic Regression, and Extra tree are the classification techniques used for this project.

E. Performance Evaluation

Accuracy, Precision, Recall, and F1-score are the four performance metrics that have been used to estimate the effectiveness of the investigated classification algorithms.Every one of the performance indicators has a specific mathematical formulation, which is described below.

Accuracy is the percentage of profiles that were successfully identified as false divided by the total number of profiles, and it reflects how close a forecast is to the actual figure.

Accuracy = (T P + T N) / (T P + T N + F P + F N) [1]

Precision: Determines the percentage of projected phoney profiles that are really positive, based on the accuracy of the categorization system. This value may be calculated using Equation 2.

Precision = TP/ (TP+FP) [2]

Recall: How many right hits were remembered (discovered), or how many genuine positives were retrieved.

Recall = TP / (TP+FN) [3]

F1-Score: When calculating the F1-score, both accuracy and recall are included in. Maximum

value it may return is 1, minimum value it can return is 0.

F1-score = (0. 2 ∗ precision ∗ recall)/ (precision + recall) [4]

RESULTS ANALYSIS

As shown in figure 2, the dataset contains 576 rows and 12 columns.

	profile pic	nums/length username	fullname words	nums/length fullname	name==username	description length	external URL	private	#posts	#followers	#follows	fake
0	1	0.27	0	0.00	0	53	0	0	32	1000	955	0
1	1	0.00	2	0.00	0	44	0	0	286	2740	533	0
2	1	0.10	2	0.00	0	0	0	1	13	159	98	0
3	1	0.00	1	0.00	0	82	0	0	679	414	651	0
4	1	0.00	2	0.00	0	0	0	1	6	151	126	0
...	...	...	...	...	...	...	...	...	...	...	...	...
571	1	0.55	1	0.44	0	0	0	0	33	166	596	1
572	1	0.38	1	0.33	0	21	0	0	44	66	75	1
573	1	0.57	2	0.00	0	0	0	0	4	96	339	1
574	1	0.57	1	0.00	0	11	0	0	0	57	73	1
575	1	0.27	1	0.00	0	0	0	0	2	150	487	1

576 rows × 12 columns

Fig. 2. Data Reading

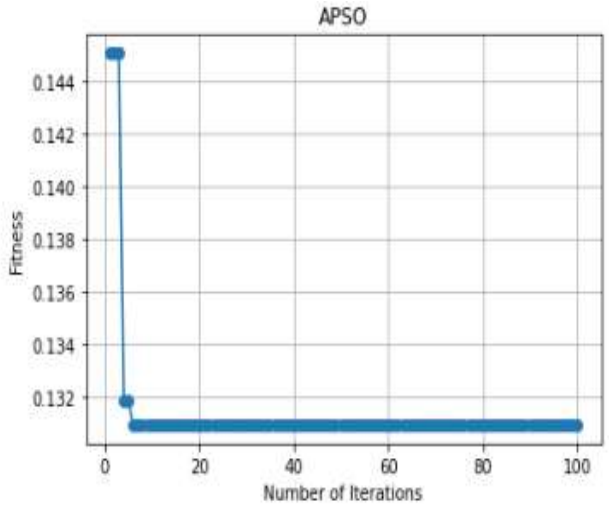

Fig. 3. APSO Feature Selectiions

TABLE II. ANALYSIS OF DIFFERENT ML ALGORITHM

Model	Accuracy			Precision			Recall			F1		
	Without	RFE	APSO	Without	RFE	APSO	Without	RFE	APSO	Without	RFE	APSO
KNN	84%	85%	96%	84%	85%	96%	84%	85%	96%	84%	85%	96%
RF	92%	85%	96%	92%	85%	96%	92%	85%	96%	92%	85%	96%
ET	90%	90%	94%	90%	90%	93%	90%	90%	94%	90%	90%	94%

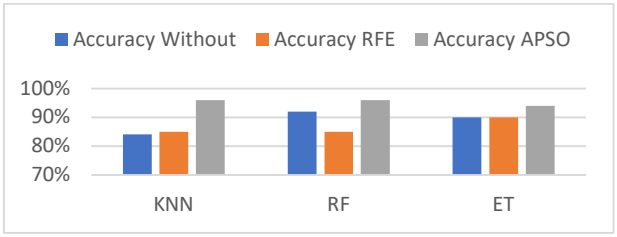

Fig. 4. Accuracy Comparative Plot

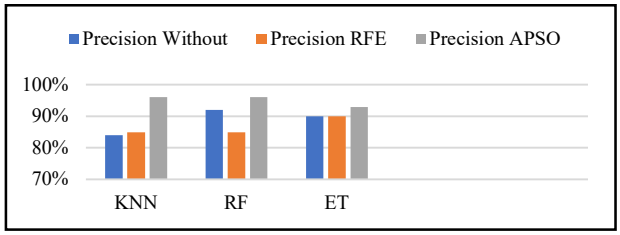

Fig. 5. Precision Comparative Plot

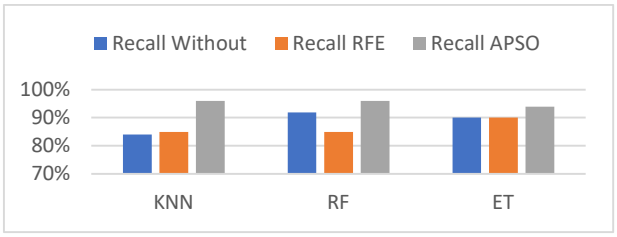

Fig. 6. Recall Comparative Plot

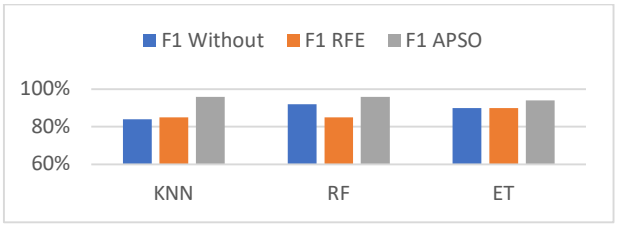

Fig. 7. F1-Score Comparative Plot

As shown in table 2, KNN, Random forest and Extra tree works better with APSO as compare with RFE and with 96% of accuracy.

CONCLUSION

Fake social networking site profiles are simple to spot using machine learning algorithms. This study proposes a revolutionary machine learning method to recognize phoney Instagram profiles using feature selection method. This research employs a two-stage procedure, first selecting features, and then classifying them. APSO, the original feature selection approach, is used to choose the useful collection of features. And machine learning algorithm, KNN,Random forest and extra tree are then applied to determine which accounts are fake and which are real. The classification algorithm's performance is heavily influenced by the choices

made when choosing its parameters. Thus, KNN and Random Forest has demonstrated its ability to accurately and appropriately identify false profiles.

REFERENCES

[1] Majed Alrubaian, Muhammad Al-Qurishi, Mohammad Mehedi Hassan, and Atif Alamri, A Credibility Analysis System for Assessing Information on Twitter, IEEE (Aug 2018) https://ieeexplore.ieee.org/document/7551232

[2] Van Der Walt, Jan Eloff and Estée. Using machine learning to detect fake identities: bots vs humans. IEEE (2018) https://ieeexplore.ieee.org/document/8265147

[3] Tri HadiahMuliawati, Perdana, Rizal Setya, and Reddy Alexandro. Bot spammer detection in Twitter using tweet similarity and time interval entropy. (Research Gate)

[4] Karatas, Arzum, and SerapŞahin. A Review on Social Bot Detection Techniques and Research Directions. (2015)

[5] Senthil Raja, M., & Arun Raj, L. (2021). Detection of malicious profiles and protecting users in online social networks. Wireless Personal Communications, 1-18.

[6] Bharti, K. K., & Pandey, S. (2021). Fake account detection in twitter using logistic regression with particle swarm optimization. Soft Computing, 25(16), 11333-11345.

[7] 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

[8] Uppada, S. K., Manasa, K., Vidhathri, B., Harini, R., &Sivaselvan, B. (2022). Novel approaches to fake news and fake account detection in OSNs: user social engagement and visual content centric model. Social Network Analysis and Mining, 12(1), 1-19.

[9] Wu, Y., Fang, Y., Shang, S., Jin, J., Wei, L., & Wang, H. (2021). A novel framework for detecting social bots with deep neural networks and active learning. Knowledge-Based Systems, 211, 106525.

[10] Sahoo, S. R., & Gupta, B. B. (2021). Real-time detection of fake account in twitter using machine-learning approach. In Advances in computational intelligence and communication technology (pp. 149-159). Springer, Singapore.

[11] Rathore, P., Soni, J., Prabakar, N., Palaniswami, M., & Santi, P. (2021). Identifying groups of fake reviewers using a semisupervised approach. IEEE Transactions on Computational Social Systems, 8(6), 1369-1378.

[12] Fang, Y., Wang, H., Zhao, L., Yu, F., & Wang, C. (2020). Dynamic knowledge graph based fake-review detection. Applied Intelligence, 50(12), 4281-429.

[13] Preethi Harris; J Gojal; R Chitra; S Anithra," Fake Instagram Profile Identification and Classification using Machine Learning", 2021 2nd Global Conference for Advancement in Technology (GCAT) | 978-1-6654-1836-2/21/$31.00 ©2021 IEEE | DOI: 10.1109/GCAT52182.2021.9587858

[14] Muñoz, S. D., & Pinto, E. P. G. (2020, December). A dataset for the detection of fake profiles on social networking services. In 2020 International Conference on Computational Science and Computational Intelligence (CSCI) (pp. 230-237). IEEE.

[15] Chaudhary, A., Mittal, H., & Arora, A. (2019, February). Anomaly detection using graph neural networks. In 2019 International Conference on Machine Learning, Big Data, Cloud and Parallel Computing (COMITCon) (pp. 346-350). IEEE.

[16] Khaled, S., El-Tazi, N., & Mokhtar, H. M. (2018, December). Detecting fake accounts on social media. In 2018 IEEE international conference on big data (big data) (pp. 3672-3681). IEEE.

[17] Chakraborty, P., Shazan, M. M., Nahid, M., Ahmed, M. K., &Talukder, P. C. (2022). Fake Profile Detection Using Machine Learning Techniques. Journal of Computer and Communications, 10(10), 74-87.

[18] Laleh, N., Carminati, B., & Ferrari, E. (2016). Risk assessment in social networks based on user anomalous behaviors. IEEE Transactions on Dependable and Secure Computing, 15(2), 295-308.

[19] Latha, P., Sumitra, V., Sasikala, V., Arunarasi, J., Rajini, A. R., &Nithiya, N. (2022, March). Fake Profile Identification in Social Network using Machine Learning and NLP. In 2022 International Conference on Communication, Computing, and Internet of Things (IC3IoT) (pp. 1-4). IEEE.