

# A Potent Technique for Identifying Fake Accounts on Social Platforms

Kajal<sup>1</sup>, Uttam Kumar Singh<sup>2</sup>, Dr. Nikhat Akhtar<sup>3</sup>, Satendra Kumar Vishwakarma<sup>4</sup>, Niranjan Kumar<sup>5</sup>, Dr. Yusuf Perweje<sup>6</sup>

<sup>1</sup>Assistant Professor, Department of Computer Science & Engineering, M.G. Institute of Management & Technology, Lucknow, India

<sup>2</sup>Assistant Professor, Department of Computer Science & Engineering, Babu Banarasi Das Northern India Institute of Technology (BBDNIIT), Lucknow, U.P, India

<sup>3</sup>Associate Professor, Department of Computer Science & Engineering, Ambalika Institute of Management & Technology, Lucknow, India

<sup>4</sup>Assistant Professor, Department of Computer Science & Engineering, Babu Banarasi Das Northern India Institute of Technology (BBDNIIT), Lucknow, U.P, India

<sup>5</sup>HoD (IT), Department of Information Technology, Ambalika Institute of Management & Technology, Lucknow, India

<sup>6</sup>Professor, Department of Computer Science & Engineering, Ambalika Institute of Management & Technology, Lucknow, U.P, India

## ARTICLE INFO

### Article History:

Accepted: 01 Aug 2023

Published: 09 Aug 2023

### Publication Issue

Volume 9, Issue 4

July-August-2023

### Page Number

308-324

## ABSTRACT

In this generation, online social media networks are rapidly growing in popularity and becoming more and more integrated into people's daily lives. These networks are used by users to exchange movies, read news articles, market products, and more. It has been simpler to add new friends and stay in touch with them and their updates. These online social networks have been the subject of research to see how they affect people. A significant amount of a user's data may attract attackers as these networks continue to develop, and these attackers may subsequently exchange incorrect information and disseminate dangerous falsehoods. Some fraudulent accounts are used to spread false information and further political agendas, for example. Finding a fraudulent account is important. Furthermore, these social networking platforms are increasingly being used by attackers to disseminate a vast amount of fake information. As a result, based on the categorization algorithms, researchers have started to investigate efficient strategies for spotting these sorts of actions and bogus accounts. In this study, various machine learning algorithms are investigated to successfully identify a phony account. To address this issue, several machine learning algorithms are utilized in conjunction with pre-processing methods to identify bogus accounts. The identification of bogus accounts uses the classification abilities of the algorithms Nave Bayes, Artificial Neural Network, Bagged Decision Tree, Radial Basis Function (RBF), Support Vector Machines, and Random Tree. The best features are used to compare the proposed model to other benchmark techniques on the dataset. The suggested Artificial Neural Network strategy outperforms the prior employed strategies to identify phony user accounts on major online social

---

platforms, with a precision of 98.90%, when machine learning techniques are also compared.

Keywords: Social Media, Fake Followers, Detection, Classification, GitHub Repositories, Machine Learning, Artificial Neural Network (ANN), Fake Account.

---

## I. INTRODUCTION

Better communication techniques have been developed throughout human history. The 1990s saw the beginning of social networking sites as a result of one such endeavor. Although it took some time for users to become accustomed to these websites, by the early 2000s the social media platform was starting to thrive [1]. Every year, at least three new social networking sites are launched. One of the largest revolutions in history was seen by the entire globe. At the moment, everyone feels the need to be online. The most popular activity on the internet now is social media [2]. A decade later, in 2020, there will be 3.6 billion active social media users, up from 970 million in 2010.

It is anticipated that social media would continue to grow during the next five years, drawing more than 4.41 billion users by 2025. These platforms have changed the internet into a social network, making it simple for [3] users to discover old acquaintances on Facebook, get the most recent news on Twitter, look for a job on LinkedIn, and watch popular videos on YouTube with just one click. These platforms have grown to be our go-to source for news and entertainment. Additionally, social media has provided us a platform where we may express our opinions without fear [4]. As a result, we have seen several online revolutions, including the well-known me-too campaign. These platforms have also developed into new types of marketing that are cost-free in contrast to conventional means of advertising like TV, radio, billboards, etc. [5]. It is now inconceivable to envisage

our lives without social media platforms since they have integrated so deeply into our daily lives [6]. Artificial intelligence (AI)'s branch of machine learning [7] enables a system to function without having to be explicitly programmed. It is utilized in a variety of industries, including cybercrimes [9], Google automobiles [8], recommendation engines, social networking networks' friend recommendations, and shopping applications. By displacing the outdated statistical methodologies, machine learning has drastically altered how data is retrieved and processed.

In order to speed up decision-making, machine learning (ML)[10] techniques are used to automate and enhance the process of identifying bogus accounts. On the other side, bots are also being created to get around those identification systems and algorithms [11]. Thus, the conflict is perpetual. But social media has a big impact on people's lives, and plenty of individuals use it to form deeper connections [12]. Instagram is among the most well-liked social networking platforms [13]. Instagram is a free social media platform designed for online picture and video sharing. It is comparable to the majority of other social media platforms [14] in that users who register an account have a profile and news feed and may share photographs and videos via both.

Many celebrities and companies have opened profiles on Instagram in recent years; they utilize the platform to increase their fan base and revenue [15]. They as well as other well-known users frequently utilize it as a platform for promotion. It is not unusual to utilize someone's account as a profitable income when they are increasing their follower count to over 100,000 or

even millions. Many celebrities and regular individuals who amassed a sizeable following on Instagram in recent years have utilized their profiles as a platform for advertising. In addition, some people aim to build their following in order to gain greater popularity, credibility, and influence. With their guidelines based on anomalous account behavior, some expert users believe that detecting phony accounts is a simple process. However, neither analytic methods to aggregate them nor validation procedures are often matched with such criteria [16]. The majority of academic academics have mostly concentrated on spam and bot identification across different social media platforms like Facebook and Twitter, with excellent results in detecting phony accounts based on their genuine and fraudulent traits, primarily using machine-learning algorithms. The empirical analysis of OSN data is heavily influenced by big data [17] issues. Big data analysis in OSNs benefits communication medium in a number of ways [18].

## II. Background

Social media is evolving really quickly right now; these services are essential to many members of society, notably for marketing [19] campaigns and for celebrities and politicians who try to promote themselves on social media by gaining followers and admirers [20]. Therefore, false accounts formed on behalf of individuals and organizations have the potential to tarnish and destroy the reputations of these individuals and organizations, ultimately resulting in a decline in the number of their actual likes and followers. By examining certain existing traits, several detection algorithms are utilized to categorize social media profiles. For improved account classification, ML algorithms [21] are among other detection methods. In [22] an Automated Feature-based fake profile detection algorithm is introduced that depends on machine learning considerations, and in [23] a new approach is introduced to detect profile cloning based on profile's attributes similarity and friend network

similarity. Other approaches are introduced to detect fake profiles based on profiles' features and behaviors. The research by Zhang et al.[24], which examined the manipulation of Twitter trends while observed from within the safety of Twitter trending, is crucial to this. This analysis uses information from more than 69% million tweets sent by 5,000,000 users. The analysis of the data reveals evidence of the use of phony and hacked accounts as vantage points to manipulate Twitter trends. A alternative method was taken by Alshaikh et al. [25], who investigated Twitter privacy and security and reported what they discovered. According to the view of Twitter users, the study, which used almost two million pertinent tweets over the period of two years, shows that there is a significant influence on privacy and security. A technique for identifying bogus accounts from the Twitter dataset is presented by Ersahin et al. [26] using the Naive Bayes classification algorithm. Entropy minimization discretization (EMD), a supervised discretization method, was used to boost the accuracy of the pre-processed dataset, yielding a 90.9% correct result. A machine learning pipeline was used in prior research [27] to identify bogus accounts on online social networks. Instead of making a forecast for each false account to see if it was made by the same person, the framework categorized groupings of bogus accounts.

Numerous categorization techniques, including support vector machine (SVM) [28], random forest, and deep neural network [29], have been presented. The growth of fake accounts is related to the popularity of social media platforms. There are many evil motives behind creating such a fake account or identity [30]. The usage of these fictitious identities is especially harmful to society since it can encourage involvement in a variety of online and offline crimes [31]. In study [32], the author developed and used a five-step automated approach for identifying spammers and harmful individuals on social media. The authors of [33] explained how to identify your close buddies in an OSN by using the Exclusive Shared Knowledge

technique. Based on the underlying behavioral characteristics of Click-Stream models, a novel technique to Sybil identification is put forward. The methodology is confirmed using ground truth traces of 16,000 actual and Sybil users from the Renren social network. In this study, the authors investigated the viability of outsourcing the Sybil detection to online human experts and they assessed the strategy on three OSNs: Renren, Facebook US, and Facebook India. Crowdsourcing [34] is a standalone approach for identifying Sybil accounts in OSNs. Similar to this, [35] provides a deep neural network-based approach that recognizes the fake profiles with the area properties by utilizing text and user information. The method [36] considers how similar friends are across different accounts in order to spot phony profiles.

A two-layer technique is also suggested for classifying profiles based on topological data and meta data [37]. Ritter et al.'s [38] weekly supervised extraction of computer security events from Twitter uses a supervised method, making it simple to define and train extractors for new categories of events by providing a limited number of seed instances. This strategy makes use of a supervised technique, which makes it simple to create and train extractors for new categories of events. Despite the fact that Twitter has not yet provided a solution to this problem, Citlak et al.'s [39] investigation highlighted how spam may be detected on the social media platform. This is plausible given that Twitter hasn't offered a solution to the problem. Comparable to this, Basha et al. [40] propose a study on developing an algorithm crucial for supplying security on Twitter by identifying the crucial holes in the chain of security occurrences. It would be required to use this algorithm to provide security on Twitter. The newly suggested technique provides higher degrees of safety compared to the two existing algorithms. Deep entity classification: Abusive account identification for online social networks was proposed by Xu, et al. [41].

Nevertheless, a realistic, successful ML-based defense has to be carefully designed with characteristics that are resistant to adversarial manipulation, enough ground truth labelled data to train the models, and a system that can scale to all active users on an OSN (perhaps billions of them). They provide Deep Entity Classification (DEC), an ML [42] framework that identifies abusive accounts in OSNs that have eluded earlier, more conventional abuse detection systems, in order to overcome these difficulties. Profiling Fake News Spreaders on social media via Psychological and Motivational Factors was suggested by Karami, et al. [43]. Most strategies created to tackle misinformation either concentrate on the fake news itself or the bad actors who produce it. Despite this, the people who spread bogus news are mostly responsible for its virality. A methodology for detecting people who are prone to distribute false news may be created with the help of a deeper understanding of these individuals. With help from psychological theories and behavioral investigations, they investigate the traits and driving forces behind those who propagate fake news on social media. The researchers next conduct a series of studies to see if people that share false news [44] vary from regular users in any way. In a prior work, managed principal component analysis (PCA) and k-mean algorithms were used [45] to identify Twitter spam accounts to improve the primary detection of spammer classes.

Several current features were used to detect spam on social networks, and new features were introduced to boost speed. Multi-layer perceptron (MLP), support vector machine, and random forest were trained as classification methods. The random forest method, which had an accuracy of 96.30%, produced the best results. Dynamic deep learning was suggested by Wanda, et al. [47] as a method for locating aberrant nodes in online social networks. The authors provide a model to classify malicious vertices by training large features with dynamic deep learning [48] architecture and leveraging connection information from nodes.

They first provide a general function called WalkPool pooling to enhance the efficiency of their network in order to build dynamic deep learning. They are able to classify the anomalous nodes with more accuracy than traditional learning algorithms by putting the suggested model to the test [49].

### III. Fake News

In the digital age, social networking sites are a very common and effective communication tool [50]. Their expansion has drawn fraudsters who steal personal information and pass it off as their own, in addition to advertising and marketing firms. These fraudsters are skilled at duping online users into disclosing private information. On social networking platforms, several false accounts can be discovered. The majority of fraudulent [51] accounts are created using other users' data without their knowledge. Usually, imposters utilize these accounts to get sensitive data from consumers, such as credit card information, other financial details, or personal information. There are rumors that political groups and businesses are creating new phony accounts to sway media trends. These accounts are often made by automated systems that mimic human behavior in order to advance a specific viewpoint or item. These accounts frequently use hash tags and repost material to boost their exposure and popularity, and there are numerous examples of this. Also noted are bogus accounts being utilized in the names of several celebrities. These accounts give an appearance of being legitimate by using the personal details and information of a celebrity. These false accounts [52] are used by imposters to trick and con lay supporters and fans.

### IV. Challenges with Fake Accounts

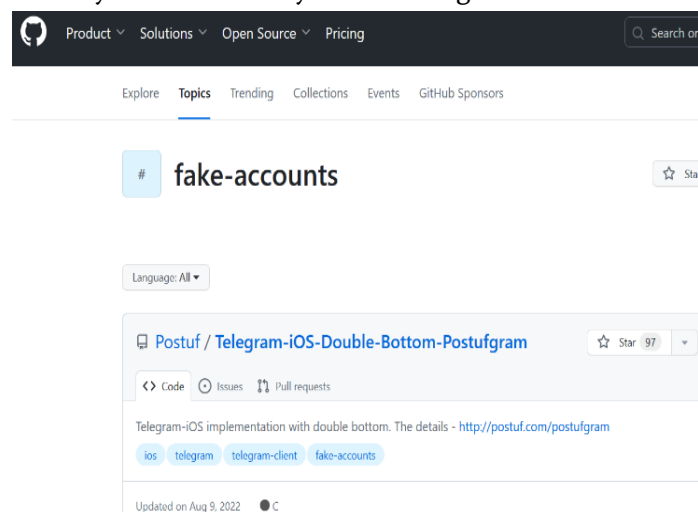
Modelling a Fake Profile Detection System is an old problem, but because of all the difficulties it poses, there are still numerous holes that have been found

and need to be filled [53]. The following is a list of the several difficulties this system poses.

- ✓ **The data is not easily accessible:** Because accounts on online social networks are so private and secure, networking websites don't disclose any account information [54] in order to preserve their confidentiality and maintain their users' confidence.
- ✓ **Both real and phony accounts share a lot of information:** When trying to train a neural network [55] to distinguish between real and false accounts, it can be difficult since the feature sets of the two types of accounts occasionally intersect.
- ✓ **The total amount of processing parameters:** The vast number of factors between learning and decision-making is a significant barrier in the development of algorithms for identifying phony accounts.
- ✓ **Choosing the best characteristics (variables) is difficult:** The choice of the best characteristics needs to be made carefully since it affects how well the system performs overall and determines which features are used to distinguish between fraudulent and actual accounts. Selecting these ideal traits might be very difficult at times.
- ✓ **The capacity to deal with data disturbances:** Disturbance is made up of missing or erroneous data, which makes it difficult to analyze the dataset. Since these systems aren't partition tolerant, there is no way to recover the lost information, which has a negative impact on the outcome.
- ✓ **A lot of the time, it looks like a normal transaction:** The activity of phony accounts [56] might occasionally closely resemble those of real ones. As a result, it becomes challenging to understand them and we stop them before they are finished.

## V. GitHub Repositories

For software engineers, GitHub is a web-based platform for version control and collaboration. The largest single donor to GitHub, Microsoft, purchased the service for \$7.5 billion in 2018. GitHub was established in 2008 and uses a software as a service (SaaS) delivery strategy. Git, an open-source code management system developed by Linus Torvalds to speed up software development, served as its foundation. Developers can modify, adapt, and enhance software from [57] GitHub's public repositories without paying a fee, but the company offers a number of premium options for private repositories. All of a project's files, together with each file's revision history, are stored in each public and private repository. Repositories can be either public or private, and they can have several collaborators. Developer cooperation is made easier by GitHub. Distributed version control is another feature offered. on order to keep organized, development teams may collaborate on a single Git repository and track changes as they happen. Along with the well-known SaaS solution, GitHub also provides an on-premises edition. GitHub Enterprise offers several third-party apps and services in addition to integrated development environments and continuous integration tools. Compared to the SaaS version, it provides more security and auditability shown in figure 1.



**Figure 1. The GitHub Repositories**

The public social platforms [58] developer API, like the majority of social media sites, only offers the users' public information. Some users' activity and login details are not accessible, usually when the user has previously changed the account to private mode. This issue is viewed as a barrier to the data collecting procedure. leverage user data from crawls to resolve problems.

## VI. Machine Learning

Artificial intelligence, which is roughly defined as a machine's capacity to mimic intelligent human behavior, includes the subfield of machine learning [59]. Artificial intelligence (AI) systems are used to carry out complicated tasks in a manner akin to how people solve issues. AI aims to develop computer models with "intelligent behaviors" similar to those of people.

This refers to devices that can comprehend a text written in natural language, identify a visual scene, or carry out a physical activity. Machine learning is one of the most exciting technologies ever created. Machine learning has a vast array of uses that address problems and automate many different sectors. The main causes of this are the expansion of informational resources, the development of machine learning [60] methodologies, and improvements in computing power. Unquestionably, machine learning has been used to a broad variety of complex and modern network administration and operation problems. For particular networking technology or specialist networking enterprises, several machine learning research have been conducted. A data [61] analysis technique called machine learning automates the creation of analytical models. Data filtering and inference are made possible by machine learning. It entails applying knowledge and advancing it through time, going beyond merely learning or receiving information [62]. The main goal of machine learning is to find and use hidden patterns in "training" data. New

data can be categorized or matched to existing categories using the patterns found [63]. The field of machine learning, which includes all branches of artificial intelligence, calls for a multidisciplinary approach that integrates knowledge from several fields, including probability theory, mathematics, trends detection, dynamic modelling (DM), cognitive psychology, adaptive control, computational neuroscience, and theoretical computer science.

### 6.1 Radial Basis Function (RBF)

An input layer, a hidden layer, and an output layer make up the three layers of feed-forward neural networks that belong to the particular class of radial basis functions. This differs significantly from the majority of neural network topologies, which have several layers and generate nonlinearity by repeatedly using nonlinear activation functions. The hidden layer is where computation takes place after receiving input data from the input layer. The Radial Basis Functions' covert layer The most potent neural network is also significantly distinct from other neural networks. The value of a radial basis function is a real-valued function whose dependence on the origin's distance is the sole other factor. The Gaussian function is the most often used radial basis function, however there are other versions as well.

### 6.2 Support Vector Machine (SVM)

A strong machine learning technique called Support Vector Machine (SVM) is utilized for applications including regression, outlier identification, and linear or nonlinear classification. Support vector machine classification aims to discriminate between two groups by giving relevant data with a feature and developing a classifier that shines on hidden data. The maximum range classification subclass of support vector machines is the most fundamental. Finding the ideal hyperplane in an N-dimensional space that may divide the data points into various classes in the feature space is the

fundamental goal of the SVM method [64]. The hyperplane strives for the widest feasible gap between the nearest points of various classes. The number of features determines the hyperplane's size. The hyperplane is essentially a line if there are just two input characteristics. The hyperplane turns into a 2-D plane if there are three input characteristics. Imagining something with more than three characteristics gets challenging.

### 6.3 Random Forest

With the aid of several decision trees and a method known as Bootstrap and Aggregation, sometimes referred to as bagging, Random Forest is an ensemble methodology capable of handling both regression and classification tasks. This method's fundamental principle is to integrate several decision trees to get the final result rather than depending just on one decision tree. Multiple decision trees serve as the fundamental learning models in Random Forest. We create sample datasets for each model by randomly selecting rows and features from the dataset. This component is known as Bootstrap.

### 6.4 Artificial Neural Network (ANN)

Artificial neurons, also known as units, are found in artificial neural networks. The entire Artificial Neural Network of a system is made up of these units, which are stacked in a number of layers. It depends on how the complicated neural networks will be used to discover the hidden patterns in the dataset whether a layer has a dozen units or millions of units [65]. Artificial neural networks frequently have hidden layers in addition to input, output, and output layers. The input layer gets information that the neural network needs to analyze or learn from the outside environment. Then, after passing through one or more hidden layers, this data is transformed into useful information for the output layer. Last but not least, the

output layer delivers an output in the form of an artificial neural network's reaction to incoming data.

### 6.5 Naive Bayes (NB)

The naive Bayes classifier assumes that each feature contributes equally and independently to the target class and is a probabilistic classifier based on the Bayes theorem [66]. According to the NB classifier, each feature independently and equally impacts the chance that a sample belongs to a certain class since it is assumed that each feature is distinct from the others and does not interact. The NB classifier is quick to implement and performs well on large datasets with high dimensionality. The NB classifier is suitable for real-time applications and is noise-resistant.

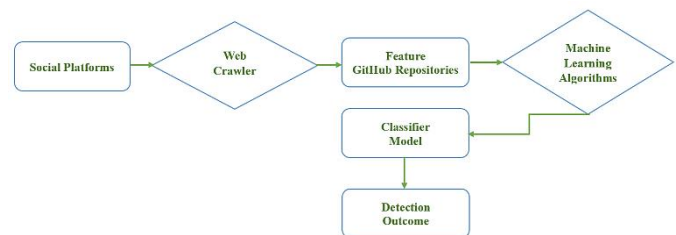
### 6.6 Bagged Decision Tree

Prediction precision is crucial for creating sound and trustworthy models in machine learning. A supervised machine-learning approach called ensemble learning integrates many models to create a more potent and reliable model. The theory is that we may build a model that is more reliable and less prone to overfit the data by integrating the advantages of many models. Both classification and regression tasks may be accomplished with it. A kind of ensemble learning known as bagging (or Bootstrap aggregating) involves the independent concurrent training of many base models on various subsets of the training data. Using bootstrap sampling, each subset is created by randomly selecting and replacing data points. For the Bagging classifier [67], the final prediction is created by combining the all-base model predictions and utilizing majority voting. Regression involves averaging all-base model predictions to arrive at the final prediction; this process is known as bagging regression.

## VII. Proposed Approach

We have provided an experimental model to construct the crucial machine learning model for the early-stage fraudulent accounts identification technique. In this part, the recommended model's specs are laid down. Using the offered dataset of fictitious accounts, we present supervised learning [68] approach assessment. The architecture of the proposed model for classifying fake news is shown in figure 2 of the methodology section. The ANN approach may be used for a variety of classification problems. By lowering the variation related to the prediction, it enhances the prediction process. The goal of the ANNs method is to calculate a number of different classifiers on gathered datasets while disrupting the training set.

Therefore, the ANNs algorithm was employed as the foundation of the suggested technique in this study to better accomplish fake users' categorization.



**Figure 2. The Architecture of Fake Accounts Detection Model**

The approach uses a crawling tool to gather user information, which is then transformed into a number of features. The ANNs algorithm used these gathered characteristics as input. After creating the dataset, we used an ANNs approach to create a classification model to determine if a certain user belongs to the real or false groups. It can differentiate between aberrant behavior and legitimate behavior by examining user behavior and content attributes since fake and genuine users exhibit distinct social behaviours.



## 7.1 Dataset

We received actual accounts from our friends' social media accounts throughout the data collecting stage. The dataset contains 7834 normal users, including famous people, businesses, and regular, lawful users, as well as 4116 anomalous individuals who were manually vetted and chosen.

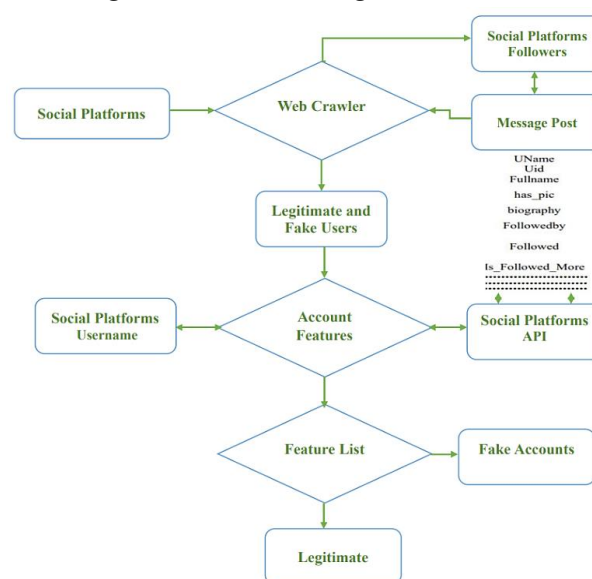
**Table 1. The List of Collected Features**

Index	Feature	Description
1	UName	Username Length
2	Uid	Real ID of user on Instagram
3	Fullname	Full name Length
4	has_pic	Does account set a profile picture
5	biography	Biography Length
6	Followedby	The number of users Followed the account
7	Followed	The number of users the account Followed them
8	Is_Followed_More	Is number of Followed are more than Followed by
9	Postcount	The number of shared posts by the account
10	is_business	Is it a business account
11	is_private	Is the user set profile as private
12	is_verified	Is the account verified by Instagram
13	has_channel	Does the account have a channel
14	external_url	Is the account linked to an external URL
15	highlight_reel_count	The number of highlights is pinned to the account
16	connected_fb_page	Is the account linked to a Facebook profile

Figure 3 illustrates how we created two different data crawler kinds for ordinary and anomalous users. The social platform explore feature was utilized by the ordinary user crawler to find normal people and add them to the dataset's list of ordinary users. Social media networks explore features highlight recently submitted images and videos that caught the attention of other users, demonstrating that the accounts posted are largely authentic and trustworthy. Furthermore, the built crawler utilized to collect fake users' IDs through the follower lists of individuals [69] who regarded a sizable number of false users in their follower list in order to locate and achieve phony [70] users on social platforms. In order to personally verify every single one of the fictitious archived users in the dataset, be certain of their identity, and raise the dataset's quality, we have created a second tool. The

following table 1 lists the characteristics that were gathered.

At last, we have examined the experimental outcomes obtained using the suggested methodology. These findings showed that the newly developed technique performs superbly in terms of accuracy and has a strong ability to identify bogus accounts. The experimental system shown in figure 2 was created in order to accurately identify GitHub Repositories and learn about data trends. The Radial Basis Function (RBF), Support Vector Machines (SVM), Random Tree, Artificial Neural Network [72], and Bagged Decision Tree algorithms are used to train the system. These models accurately build trained models by accounting for the training GitHub Repositories. Once these models are learned, they can be used to categorize data using the four test GitHub repositories that were created [57] and yield accurate results for the test datasets. The results of this model's performance are shown in table 2 below. When it came to performance metrics like classification accuracy for the task of identifying bogus accounts, we also compared the suggested Machine Learning classifier to the benchmark models. The dataset that was created included news items, of which about 70% were used for training and 30% for testing.



**Figure 3. Dataset and Feature Collection Process**

### VIII. Results Assessment

In this part, several comparisons and the usefulness of the machine learning algorithm in identifying bogus news are addressed. For their performance studies, certain performance metrics are measured [71]. In all of our studies, we have also used the gathered dataset, which contains data on 10,000 users of social media sites. These social user samples fall into the two primary categories of normal and abnormality.

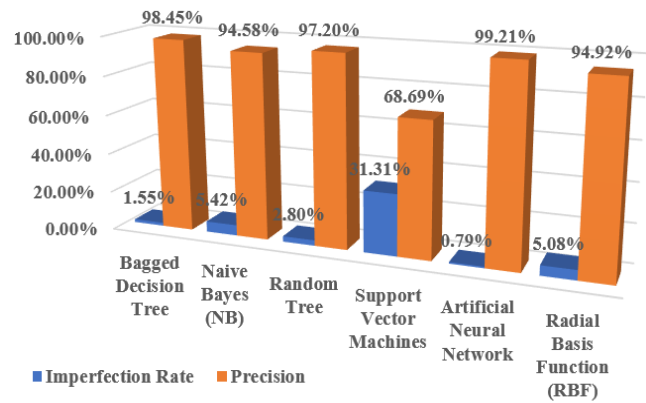
**Table 2. The Model's Performance of Classification Outcomes for the Test Datasets**

Machine Learning Algorithms	Performance Summary for 70% - 30%	
	Imperfection Rate	Precision
Bagged Decision Tree	1.55%	98.45%
Naive Bayes (NB)	5.42%	94.58%
Random Tree	2.80%	97.20%
Support Vector Machines	31.31%	68.69%
Artificial Neural Network	0.79%	99.21%
Radial Basis Function (RBF)	5.08%	94.92%
Bagged Decision Tree	1.55%	98.45%

Figure 4 displays the algorithms' accuracy for the overall precision. The algorithm's accuracy is represented in this case as a percentage (%). The results show that the algorithm performs well, with a ratio of 70 to 30 [73]. Performance is measured by the algorithm's imperfection rate, which shows how frequently [74] the algorithm is misclassified. This equation might be used to compute that.

$$\text{Imperfection Rate} = 100 - \text{Precision}$$

In addition, we discovered that ANNs perform better than the other used methods (see figure 4). Therefore, strategies may be considered for the implementation of the suggested data model in the near future.



**Figure 4. The Model's Performance Summary**

### 8.1 Performance Measurements

For the purpose of identifying fake accounts, the suggested approach was given a sizable number of news articles as input. After retrieving the classification outcomes in terms of the confusion matrix, the performance metrics accuracy (A), precision (P), recall (R), and F-measure (F) were evaluated using the following equations.

$$\text{Accuracy (A)} = \frac{T_p + T_n}{T_p + F_p + T_n + F_n}$$

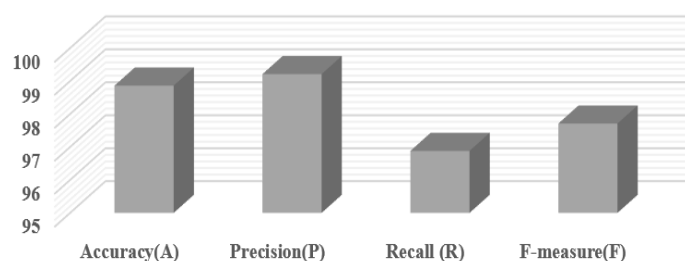
$$\text{Precision (P)} = \frac{T_p}{T_p + F_p}$$

$$\text{Recall (R)} = \frac{T_p}{T_p + F_n}$$

$$\text{F-Measure (F)} = \frac{2 * P * R}{P + R}$$

The quantity of news items accurately identified as unfavorable for a particular news category,  $T_n$ , is the true negative value.  $T_p$  is the number of news items accurately classified [75] as positive for a given news category. The number of news items that should not be identified as falling within the specified category but are yet labelled as such is known as the false positive value ( $F_p$ ). The number of items that should not be categorized as belonging to the given category but are

yet classified as such is known as the false negative value (Fn). Several query articles were submitted to the system based on these parameters, and the A,P,R, and F values were recorded. Table 2 provides a summary of this finding across different article types and evaluates the average accuracy values for the ANNs [76] models.



**Figure 5. The ANNs Model Performance Metrics Accuracy(A), Precision(P), Recall(R), and F-measure(F)**

## IX. Conclusion

Social media platforms like Instagram, Facebook, Twitter, and others are used by people today for a variety of things, including collecting knowledge, purchasing, promoting, and selling consultancy, instruction, community development, chatting with their pals while having fun. For social media platforms progressively cover a due to this. they play a significant role in our everyday lives and are influential is expanding throughout our life. The concern of certain accounts manipulating and disseminating false information on social media platforms has grown significantly. In this research, a machine learning-based technique for detecting bogus accounts on social media platforms is presented. In order to determine if provided social network account credentials are from real or bogus users, this study employs an ANN. We have built a dataset of real and fraudulent Instagram profiles in order to achieve the objective of the suggested strategy. When fresh test data is provided for prediction, the trained model will be used to the new test data to determine if the given new account information are from legitimate or false users. The ANN algorithm will be trained with all previous users'

fake and genuine account datasets. The newly developed method for distinguishing between fraudulent and real accounts took into consideration the user's content and behavior characteristics and applied them to the bagging classifier algorithm. As a consequence, after extensive research, experimentation, assessment, and implementation work, the results of the studies have demonstrated that the suggested approach is workable and capable of properly identifying over 99% of users.

## Data Availability

The study used open-source dataset and is accessed from the Weblink <https://github.com/harshitkgupta/Fake-Profile-Detection-using-ML>

## X. REFERENCES

- [1]. Z. Chu, S. Gianvecchio, H. Wang and S. Jajodia, "Detecting automation of twitter accounts: Are you a human bot or cyborgf", IEEE Transactions on Dependable and Secure Computing, vol. 9, no. 6, pp. 811-824, 2012
- [2]. Edosomwan, Simeon & Prakasan, S.K. & Kouame, D. & Watson, J. & Seymour, T.. (2011). The history of social media and its impact on business. Journal of Applied Management and Entrepreneurship. 16. 79-91.
- [3]. Suja P Mathews, Sunu George, "Growth and Future of Social Media", International Journal of Advanced Research in Computer Engineering & Technology (IJARCET) Volume 2, Issue 12, December 2013
- [4]. Vishal Verma, Apoorva Dwivedi, Kajal, Prof. (Dr.) Devendra Agarwal, Dr. Fokrul Alom Mazarbhuiya, Dr. Yusuf Perwej, "An Evolutionary Fake News Detection Based on Tropical Convolutional Neural Networks (TCNNs) Approach", International Journal of Scientific Research in Science and Technology

- (IJSRST), Print ISSN: 2395-6011, Online ISSN: 2395-602X, Volume 10, Issue 4, Pages 266-286, July-August-2023, DOI: 10.32628/IJSRST52310421
- [5]. Adikari, Shalinda, and Kaushik Dutta. "Identifying Fake Profiles in LinkedIn." In PACIS, p. 278. 2014
- [6]. Ms Farah Shan, Versha Verma, Apoorva Dwivedi, Dr. Yusuf Perwej, Ashish Kumar Srivastava, "Novel Approaches to Detect Phony Profile on Online Social Networks (OSNs) Using Machine Learning", International Journal of Scientific Research in Computer Science, Engineering and Information Technology (IJSRCSEIT), Volume 9, Issue 3, Pages 555-568, May-June 2023, DOI: 10.32628/CSEIT23903126
- [7]. Yusuf Perwej, Firoj Parwej, "A Neuroplasticity (Brain Plasticity) Approach to Use in Artificial Neural Network", International Journal of Scientific & Engineering Research (IJSER), France , ISSN 2229 – 5518, Volume 3, Issue 6, Pages 1- 9, 2012, DOI: 10.13140/2.1.1693.2808
- [8]. Sachin Bhardwaj, Apoorva Dwivedi, Ashutosh Pandey, Dr. Yusuf Perwej, Pervez Rauf Khan, "Machine Learning-Based Crowd Behavior Analysis and Forecasting", International Journal of Scientific Research in Computer Science, Engineering and Information Technology (IJSRCSEIT), ISSN: 2456-3307, Volume 9, Issue 3, Pages 418-429, May-June 2023-2023, DOI: 10.32628/CSEIT23903104
- [9]. Yusuf Perwej, Prof. (Dr.) Syed Qamar Abbas, Jai Pratap Dixit, Nikhat Akhtar, Anurag Kumar Jaiswal, "A Systematic Literature Review on the Cyber Security", International Journal of Scientific Research and Management (IJSRM), ISSN (e): 2321-3418, Volume 9, Issue 12, Pages 669 - 710, 2021, DOI: 10.18535/ijprm/v9i12.ec04
- [10]. Venkata K. S. Maddala, Dr. Shantanu Shahi, Dr. Yusuf Perwej, H G Govardhana Reddy, "Machine Learning based IoT application to Improve the Quality and precision in Agricultural System", European Chemical Bulletin (ECB), ISSN: 2063-5346, SCOPUS, Hungary, Volume 12, Special Issue 6, Pages 1711 – 1722, May 2023, DOI: 10.31838/ecb/2023.12.si6.157
- [11]. Bhavesh Kumar Jaisawal, Dr. Yusuf Perwej, Sanjay Kumar Singh, Susheel Kumar, Jai Pratap Dixit, Niraj Kumar Singh, "An Empirical Investigation of Human Identity Verification Methods" , International Journal of Scientific Research in Science, Engineering and Technology (IJSRSET), Print ISSN: 2395-1990 , Volume 10, Issue 1, Pages 16-38, July-August-2022, DOI: 10.32628/IJSRSET2310012
- [12]. Lőrincz, L., Koltai, J., Győr, A.F., Takács, K. (2019). Collapse of an online social network: Burning social capital to create it? Social Networks, 57: 43-53. <https://doi.org/10.1016/j.socnet.2018.11.004>
- [13]. Arora, A., Bansal, S., Kandpal, C., Aswani, R., Dwivedi, Y. (2019). Measuring social media influencer index435 insights from Facebook, Twitter and Instagram. Journal of Retailing and Consumer Services, 49: 86-101.
- [14]. Apoorva Dwivedi, Dr. Basant Ballabh Dumka, Susheel Kumar, Dr. Fokrul Alom Mazarbhuiya, Ms Farah Shan, Dr. Yusuf Perwej, "State of the Art Machine Learning Techniques for Detecting Fake News", International Journal of Scientific Research in Science, Engineering and Technology (IJSRSET), Print ISSN: 2395-1990, Online ISSN: 2394-4099, Volume 10, Issue 4, Pages 115-130, July-August 2023. DOI: 10.32628/IJSRSET23103191
- [15]. Boerman, S.C. (2020). The effects of the standardized Instagram disclosure for micro-and meso-influencers. Computers in Human Behavior, 103: 199-207. <https://doi.org/10.1016/j.chb.2019.09.015>
- [16]. Saurabh Sahu, Km Divya, Dr. Neeta Rastogi, Puneet Kumar Yadav, Dr. Yusuf Perwej, "Sentimental Analysis on Web Scraping Using

- Machine Learning Method” , Journal of Information and Computational Science (JOICS), ISSN: 1548-7741, Volume 12, Issue 8, Pages 24-29, August 2022, DOI: 10.12733/JICS.2022/V12I08.535569.67004
- [17]. Yusuf Perwej, Md. Husamuddin, Fokrul Alom Mazarbhuiya, “An Extensive Investigate the MapReduce Technology”, International Journal of Computer Sciences and Engineering (IJCSE), Volume-5, Issue-10, Page No. 218-225, Oct-2017, DOI: 10.26438/ijcse/v5i10.218225
- [18]. Elyusufi, Yasyn & Elyusufi, Zakaria & M'hamed, Ait Kbir. , “Social Networks Fake Profiles Detection Using Machine Learning Algorithms”. 10.1007/978-3-030-37629-1\_3, 2020
- [19]. Prof. Kameswara Rao Poranki, Dr. Yusuf Perwej, Dr. Asif Perwej,” The Level of Customer Satisfaction related to GSM in India “, TIJ’s Research Journal of Science & IT Management – RJSITM, International Journal's-Research Journal of Science & IT Management of Singapore, ISSN: 2251-1563, Singapore, Volume 04, Number: 03, Pages 29-36, 2015
- [20]. Blair, S.J., Bi, Y., Mulvenna, M.D., “Aggregated topic models for increasing social media topic coherence”, Applied Intelligence, 50(1): 138-156, 2020
- [21]. Shweta Pandey, Rohit Agarwal, Sachin Bhardwaj, Sanjay Kumar Singh, Dr. Yusuf Perwej, Niraj Kumar Singh, “A Review of Current Perspective and Propensity in Reinforcement Learning (RL) in an Orderly Manner” , International Journal of Scientific Research in Computer Science, Engineering and Information Technology (IJSRCSEIT), Volume 9, Issue 1, Pages 206-227, January-February-2023, DOI: 10.32628/CSEIT2390147
- [22]. M. Fire, G. Katz, and Y. Elovici," Strangers Intrusion Detection Detecting Spammers and Fake Profiles in Social Networks Based on Topology Anomalies", HUMAN, 1(1), pp: 26-39, 2012
- [23]. M.R. Khayyambashi, and F.S. Rizzi, "An Approach for Detecting Profile Cloning in Online Social Networks", 7th International Conference on e-Commerce in Developing Countries with Focus on e-Security (ECDCC), IEEE, 2013
- [24]. Zhang, Y.; Ruan, X.; Wang, H. Wang, H.; He, S. “Twitter Trends Manipulation: A First Look Inside the Security of Twitter Trending”, IEEE Trans. Inf. Forensics Secur., 12, 1, 2016
- [25]. Alshaikh, M.; Zohdy, M.; Olawoyin, R.; Debnath, D.; Gwarzo, Z.; Alowibdi, J. Social Network Analysis and Mining: Privacy and Security on Twitter. In Proceedings of the 2020 10th Annual Computing and Communication Workshop and Conference (CCWC), Las Vegas, NV, USA, 6–8 January 2020; pp. 0712–0718.
- [26]. B. Ersahin, O. Aktas, D. Kilinc, and C. Akyol, “Twitter fake account detection,” in 2017 International Conference on Computer Science and Engineering (UBMK), Oct. 2017, pp. 388–392, doi: 10.1109/UBMK.2017.8093420.
- [27]. G. A, Radhika S, and Jayalakshmi S. L., “Detecting fake accounts in media application using machine learning,” International Journal of Advanced Networking and Applications (IJANA), pp. 234–237.
- [28]. Asif Perwej, Prof. (Dr.) K. P. Yadav, Prof. (Dr.) Vishal Sood, Dr. Yusuf Perwej, “ An Evolutionary Approach to Bombay Stock Exchange Prediction with Deep Learning Technique”, IOSR Journal of Business and Management (IOSR-JBM), e-ISSN: 2278-487X, p-ISSN: 2319-7668, USA, Volume 20, Issue 12, Ver. V, Pages 63-79, 2018, DOI: 10.9790/487X-2012056379
- [29]. Yusuf Perwej, “The Bidirectional Long-Short-Term Memory Neural Network based Word Retrieval for Arabic Documents”, Transactions on Machine Learning and Artificial Intelligence (TMLAI), Society for Science and Education, United Kingdom (UK), ISSN 2054-7390, Volume

- 3, Issue 1, Pages 16 - 27, 2015, DOI: 10.14738/tmlai.31.863
- [30]. Y. Boshmaf, D. Logothetis, G. Siganos, J. Lería, J. Lorenzo, M. Ripeanu, K. Beznosov, H. Halawa, "Íntegro: Leveraging victim prediction for robust fake account detection in large scale osns", *Computers & Security*, vol. 61, pp. 142-168, 2016
- [31]. Asif Perwej, Kashiful Haq, Yusuf Perwej, "Blockchain and its Influence on Market", *International Journal of Computer Science Trends and Technology (IJCST)*, ISSN 2347 – 8578, Volume 7, Issue 5, Pages 82- 91, 2019, DOI: 10.33144/23478578/IJCST-V7I5P10
- [32]. Gao, J. Hu, and C. Wilson, "Detecting and Characterizing Social Spam Campaigns "Proceedings of the 10th ACM SIGCOMM conference on Internet measurement. ACM, PP: 35-47, (2010).
- [33]. R. Baden, N. Spring, and B. Bhattacharjee, "Identifying Close Friends on the Internet", 8th ACM Workshop on Hot Topics in Networks , HotNets, New York, NY, USA,(2009).
- [34]. G. Wang, M. Mohanlal, C. Wilson, X. Wang, M. Metzger, H. Zheng, and B.Y. Zhao, "Social Turning Tests: Crowdsourcing Sybil Detection", *The 20th Network & Distributed System Security Symposium, NDSS, 2013*
- [35]. Y. Liu and Y. F. B. Wu, "Fned: a deep network for fake news early detection on social media", *ACM Transactions on Information Systems*, Vol. 38, pp. 1-33, 2020
- [36]. M. Mohammadrezaei, M. E. Shiri, and A. M. Rahmani, "Identifying fake accounts on social networks based on graph analysis and classification algorithms", *Security and Communication Networks*, 2018
- [37]. Firoj Parwej, Nikhat Akhtar, Dr. Yusuf Perwej, "A Close-Up View About Spark in Big Data Jurisdiction", *International Journal of Engineering Research and Application (IJERA)*, ISSN: 2248-9622, Volume 8, Issue 1, (Part -I1), Pages 26-41, 2018, DOI: 10.9790/9622-0801022641
- [38]. Ritter, A.; Wright, E.; Casey, W.; Mitchell, T. Weakly Supervised Extraction of Computer Security Events from Twitter. In *Proceedings of the 24th International Conf. on World Wide Web*, Geneva, Switzerland, 18–22, pp. 896–905, 2015
- [39]. Çıtlak, O.; Dörterler, M.; Doğru, I.A. A survey on detecting spam accounts on Twitter network. *Soc. Netw. Anal. Min.* 2019, 9, 1–13.
- [40]. Basha, C.B.; Rajaprakash, S.; Harish, V.V.A.; Krishna, M.S.; Prabhas, K. Securing Twitter Analysed Data Using CBB22 Algorithm. *Adv. Math. Sci. J.* 2020, 9, 1093–1099.
- [41]. Xu, Teng, et al. "Deep entity classification: Abusive account detection for online social networks." *30th {USENIX} Security Symposium ({USENIX} Security 21)*. 2021.
- [42]. Y. Perwej, Firoj Parwej, "A Neuroplasticity (Brain Plasticity) Approach to Use in Artificial Neural Network", *International Journal of Scientific & Engineering Research (IJSER)*, France , ISSN 2229 – 5518, Volume 3, Issue 6, Pages 1- 9, 2012, DOI: 10.13140/2.1.1693.2808
- [43]. Karami, Mansooreh, Tahora H. Nazer, and Huan Liu. "Profiling Fake News Spreaders on Social Media through Psychological and Motivational Factors." *Proceedings of the 32nd ACM Conference on Hypertext and Social Media*. 2021.
- [44]. S. Helmstetter and H. Paulheim, "Weakly supervised learning for fake news detection on Twitter", 2018 *IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM)*, pp. 274-277, 2018
- [45]. K. S. Adewole, T. Han, W. Wu, H. Song, and A. K. Sangaiah, "Twitter spam account detection based on clustering and classification methods," *The Journal of Supercomputing*, vol. 76, no. 7,

- pp. 4802–4837, Jul. 2020, doi: 10.1007/s11227-018-2641-x.
- [46]. Neha Kulshrestha, Nikhat Akhtar, Dr. Yusuf Perwej, “Deep Learning Models for Object Recognition and Quality Surveillance”, International Conference on Emerging Trends in IoT and Computing Technologies (ICEICT-2022), ISBN 978-10324-852-49, SCOPUS, Routledge, Taylor & Francis, CRC Press, Chapter 75, pages 508-518, Goel Institute of Technology & Management, 2022, DOI: 10.1201/9781003350057-75
- [47]. Wanda, Putra, and Huang J. Jie. "Deep Friend: finding abnormal nodes in online social networks using dynamic deep learning." *Social Network Analysis & Mining* 11.1, 1- 12, 2021
- [48]. Y. Perwej, “An Evaluation of Deep Learning Miniature Concerning in Soft Computing”, International Journal of Advanced Research in Computer and Communication Engineering (IJARCCE), ISSN (Online): 2278-1021, ISSN (Print): 2319-5940, Volume 4, Issue 2, Pages 10 - 16, 2015 , DOI: 10.17148/IJARCCE.2015.4203
- [49]. T M Cover and P E. Hart, "Nearest neighbor pattern classification", *IEEE Transactions on Information Theory*, vol. 13, no. 1, pp. 21-27, 1967
- [50]. Zhang ve H. Sun, “Instagram spam detection,” *IEEE 22nd Pacific Rim International Symposium on Dependable Computing (PRDC)*, sf. 227–228, 2017
- [51]. Y. Perwej, Kashiful Haq, Uruij Jaleel, Firoj Perwej, “Block Cipherring in KSA, A Major Breakthrough in Cryptography Analysis in Wireless Networks”, *International Transactions in Mathematical Sciences and Computer*, India, ISSN-0974-5068, Volume 2, No. 2, Pages 369-385, 2009
- [52]. Q. Cao, X. Yang, J. Yu, ve C. Palow, “Uncovering large groups of active malicious accounts in online social networks,” *Proceedings of the ACM Conference on Computer and Communications Security*, sf. 477–488, 2014
- [53]. E. Van Der Walt and J. Eloff, "Using Machine Learning to Detect Fake Identities: Bots vs Humans," *IEEE Access*, 6, 6540-6549, 2018
- [54]. Saurabh Sahu, Km Divya, Dr. Neeta Rastogi, Puneet Kumar Yadav, Dr. Yusuf Perwej, “Sentimental Analysis on Web Scraping Using Machine Learning Method” , *Journal of Information and Computational Science (JOICS)*, ISSN: 1548-7741, Volume 12, Issue 8, Pages 24-29, August 2022, DOI: 10.12733/JICS.2022/V12I08.535569.67004
- [55]. Y. Perwej, Asif Perwej, “Forecasting of Indian Rupee (INR) / US Dollar (USD) Currency Exchange Rate Using Artificial Neural Network”, *International Journal of Computer Science, Engineering and Applications (IJCSEA)*, Academy & Industry Research Collaboration Center (AIRCC), USA , Volume 2, No. 2, Pages 41- 52, 2012, DOI: 10.5121/ijcsea.2012.2204
- [56]. Kim S., Han J., “Detecting Engagement Bots on Social Influencer Marketing”, In: , et al. *Social Informatics. SocInfo 2020. Lecture Notes in Computer Science*, 12467, Springer, Cham, 2020
- [57]. <https://github.com/harshitkgupta/Fake-Profile-Detection-using-ML/commit/1507d5b906e9818b37d36f1c976acb3fcb81c6a0>
- [58]. Pratama R. P. and Tjahyanto A., “The influence of fake accounts on sentiment analysis related to COVID-19 in Indonesia”, *Procedia Computer Science*, 197, 143-150, 2022
- [59]. Y. Perwej, “An Evaluation of Deep Learning Miniature Concerning in Soft Computing”, *International Journal of Advanced Research in Computer and Communication Engineering (IJARCCE)*, ISSN (Online): 2278-1021, ISSN (Print): 2319-5940, Volume 4, Issue 2, Pages 10 - 16, 2015, DOI: 10.17148/IJARCCE.2015.4203

- [60]. Li, D. Li and C. Zhang, "An Application of Machine Learning in the Criterion Updating of Diagnosis Cancer", IEEE, pp. 187-190, 2005
- [61]. Y. Perwej, Bedine Kerim, Mohmed Sirelkhitem Adrees, Osama E. Sheta, "An Empirical Exploration of the Yarn in Big Data", International Journal of Applied Information Systems (IJ AIS) – ISSN: 2249-0868, Foundation of Computer Science FCS, New York, USA, Volume 12, No.9, Pages 19-29, December 2017, DOI: 10.5120/ijais2017451730
- [62]. A. El Azab, A. M. Idrees, M. A. Mahmoud and H. Hefny, "Fake account detection in twitter based on minimum weighted feature set", Int. Scholarly Sci. Res. Innov., vol. 10, no. 1, pp. 13-18, 2016
- [63]. Yusuf Perwej, Dr. Shaikh Abdul Hannan, Firoj Parwej, Nikhat Akhtar, "A Posteriori Perusal of Mobile Computing", International Journal of Computer Applications Technology and Research (IJCATR), ATS (Association of Technology and Science), India, ISSN 2319–8656 (Online), Volume 3, Issue 9, Pages 569 - 578, 2014, DOI: 10.7753/IJCATR0309.1008
- [64]. A. Statnikov, D. Hardin and C. Aliferis, Using SVM weight-based methods to identify causally relevant and non-causally relevant variables, 2006
- [65]. B. K. Bose, "Neural network applications in power electronics and motor drives – an introduction and perspective", IEEE Trans. Ind. Electron., vol. 54, no. 1, pp. 14-33, 2007
- [66]. Nikhat Akhtar, Devendera Agarwal, "An Efficient Mining for Recommendation System for Academics", International Journal of Recent Technology and Engineering (IJRTE), ISSN 2277-3878 (online), SCOPUS, Volume-8, Issue-5, Pages 1619-1626, 2020, DOI: 10.35940/ijrte.E5924.018520
- [67]. J. Dou, A. P. Yunus, D. T. Bui, A. Merghadi, M. Sahana, Z. Zhu, et al., "Improved landslide assessment using support vector machine with bagging boosting and stacking ensemble machine learning framework in a mountainous watershed Japan", Landslides, vol. 17, no. 3, pp. 641-658, 2019
- [68]. M. J. Patwary and X. Z. Wang, "Sensitivity analysis on initial classifier accuracy in fuzziness based semi-supervised learning", Information Sciences, vol. 490, pp. 93-112, 2019
- [69]. Firoj Parwej, Nikhat Akhtar, Dr. Yusuf Perwej, "An Empirical Analysis of Web of Things (WoT)", International Journal of Advanced Research in Computer Science (IJARCS), ISSN: 0976-5697, Volume 10, No. 3, Pages 32-40, May 2019, DOI: 10.26483/ijarcs.v10i3.6434
- [70]. Y. Perwej, Dr. Nikhat Akhtar, Neha kulshrestha, Pavan Mishra, "A Methodical Analysis of Medical Internet of Things (MIoT) Security and Privacy in Current and Future Trends", Journal of Emerging Technologies and Innovative Research (JETIR), ISSN-2349-5162, Volume 09, Issue 1, Pages 346 - 371, January 2022 , DOI: 10.6084/m9.figshare.JETIR2201346
- [71]. Davis, J., and Goodrich, M. : The relationship between Precision- Recall and ROC curves. 23rd international conference on machine learning, 233-240, 2006
- [72]. Giuseppe Fenza, Francesco Orciuoli and Demetrios G. Sampson, "Building Adaptive Tutoring Model using Artificial Neural Networks and Reinforcement Learning", IEEE 17th International Conference on Advanced Learning Technologies (ICALT), pp. 2161-377X/17, 2017
- [73]. Jesse Davis and Mark Goodrich. The relationship between precision-recall and roc curves. In Proceedings of the 23rd international conference on Machine learning, pages 233–240. ACM, 2006
- [74]. Takaya Saito and Marc Rehmsmeier. The precision-recall plot is more informative than the roc plot when evaluating binary classifiers on imbalanced datasets. PloS one, 10(3):e0118432, 2015



- [75]. Apoorva Dwivedi, Dr. Basant Ballabh Dumka, Dr. Nikhat Akhtar, Ms Farah Shan, Dr. Yusuf Perwej, "Tropical Convolutional Neural Networks (TCNNs) Based Methods for Breast Cancer Diagnosis", International Journal of Scientific Research in Science and Technology (IJSRST), Print ISSN: 2395-6011, Online ISSN: 2395-602X, Volume 10, Issue 3, Pages 1100 - 1116, May-June-2023, DOI: 10.32628/IJSRST523103183
- [76]. J.-W. Lin and J.-S. Chiou, "Active probability backpropagation neural network model for monthly prediction of probabilistic seismic hazard analysis in Taiwan", IEEE Access, vol. 7, pp. 108990-109014, 2019

**Cite this article as :**

Kajal, Uttam Kumar Singh, Dr. Nikhat Akhtar, Satendra Kumar Vishwakarma, Niranjan Kumar, Dr. Yusuf Perwej, "A Potent Technique for Identifying Fake Accounts on Social Platforms", International Journal of Scientific Research in Computer Science, Engineering and Information Technology (IJSRCSEIT), ISSN : 2456-3307, Volume 9, Issue 4, pp.308-324, July-August-2023. Available at doi : <https://doi.org/10.32628/CSEIT2390425>  
Journal URL : <https://ijsrcseit.com/CSEIT2390425>