Detection of Fake Twitter Accounts using Ensemble Learning Mode

Suprith KS
Dept. Of Computer Engineering
Ramrao Adik Institute Of Technology
Navi Mumbai, India
suprith.ks01@gmail.com

Nikhil Jamdar
Dept. Of Computer Engineering
Ramrao Adik Institute Of Technology
Navi Mumbai, India
nikhiljamdar101@gmail.com

Jai Gangan
Dept. Of Computer Engineering
Ramrao Adik Institute Of Technology
Navi Mumbai, India
jaigangan10@gmail.com

Smita Bharne
Dept. Of Computer Engineering
Ramrao Adik Institute Of Technology
Navi Mumbai, India
smita.bharne@rait.ac.in

***Abstract*—The main purpose of our fake account detection is to eliminate all fake accounts that can detect if the account is fake before the user takes any action on the network. Advanced attacks and other threats are more effective against counterfeiting. The number of Internet Behavior Information of consumers uses communication to perform daily tasks such as cloud sharing, reading news, sending messages, reviewing products, chatting about events. In addition, other scammers are attracted to these social media channels. These cybercriminals include trolls, internet fraudsters, advertising campaigns, and sexual predators, among others. To distribute their content and commit fraud, these guys are setting up phone accounts. Both users and service providers are seriously harmed by all of the malicious identities. To identify those accounts and determine if they are real or fake, go to the social media service providers. In this research, we put forth a variety of categorization algorithms, including the Naive Bayes and Decision Tree algorithms. This algorithm aids in identifying fake social media profiles.**

Keywords—fake profile, Twitter, online social network, cybercrime, ensemble learning model

I. INTRODUCTION

Social media is used and run by each and every growing individual, company, business house, educational or other institutions, etc. Social media platforms and its parts and parcels are influencing people to such an extent, that in some cases, it's not the individuals running social media, but social media running them.[1] In this age where social media and its various platforms are ruling the big industries and aspects of work and economy, as well as social and interpersonal lives, there has to be an efficient check on the setbacks, inappropriate and negative aspects as well as the problems and bugs in the entire system. While this whole social media facet and semblance is at an all-time and ever expanding high of usage and influence, there are its own names which give rise to issues like privacy issues, mishandling of liberty provided under the controllable social media platforms, issue of bots, ingredients of fake propaganda and so on.[8] As the issue of fake account creation and dishonest misappropriation and impersonation is elevating, systems to keep a check on the same and classify the real and fake aspects would come to great help. The assignment worked on and here, he presents a machine learning-based solution for identifying bogus Twitter accounts. A number of advantages can be derived through this including detection of fake and real accounts to an extent and thereby avoiding fraudulent impersonation or misappropriation, making the social media platforms more secure for use and the list goes on[5].

II. LITERATURE SURVEY

The systems for detecting bogus accounts have been built by a number of authors. The most current innovations proposed through papers are outlined below:
In paper [1]'Fake Profile Identification using Machine Learning Algorithms' - Mamatha Mallam Peta and her team provide a study of the various data mining techniques employed to find anomalies. Analyzing informal community-driven criminal justice systems, commonly known as crime-based, pattern-based, and fictitious-based systems, requires a different approach. Every single person in this group also combines a number of the processes discussed in the study. The report has been concluded with a number of potential future themes and research areas that could be attended to and pursued.

Inpaper[2]'Detecting fake account on social media using machine learning algorithms' Twitter trends are safe from manipulation by haters. From 5 million records, we collect more than 69 million tweets. To find proof of Twitter pattern control, we first undertake an information investigation using the collected tweets. At that point, we learn about the particular subject. and identify the critical criteria that determine whether a theme begins to skew due to its predominance, inclusion, transmission, potential in collusion, or renown. We discover that, with the exception of transmission, each of the aforementioned components is clearly associated with inclining. Finally, we examine slope control for fraud and information exchange.

In article"Exploring Clusters of Fake Accounts on Online Social Networks" [3] Cao Xiao described a scalable method for finding groups of fake accounts created by humans. The simple way is to monitor the machine learning pipeline to determine if the entire financial system is bad. The statistics of the user who created the text, such as name, email, workplace or school, anything that indicates the frequency of the

sample in the group, are the main characteristics used in the sample. LinkedIn account information generated by registration IP address and registration date will be analyzed using the framework.

In article [4] "Fake Profile Identification Using Machine Learning" - In order to stop the propagation of incorrect or damaging information, Samrala Durga Prasad Reddy reveals that they have created a new trust method to gauge Twitter gender right to information. The feature sorting algorithm, the knowledge-based component, the performance rating, the user information component, and the knowledge-based component are built into the framework in four different ways. Together, these elements form an algorithmic model that assesses the legitimacy of Twitter tweets and users. They also assessed the operation's effectiveness using two separate databases of the top 489,330 Twitter accounts.

In the paper [5] 'Fake Account Detection Using Machine Learning'- used boosting methods rather than the usual machine learning classifiers to increase the accuracy of the standard method. This approach has significantly increased accuracy by strengthening poor learners. The accuracy of the Xgboost Classifier and the Gradient Boosting Classifier are compared by T. Om Pratyahara and team. In paper [12] author has used the various classification algorithm such as support vector machine, random forest, decision tree classifier to detect the fake account from online social network. Using ensemstack algorithm they got accuracy of 94.20%.

The existing systems are listed in Table 2.1. We looked into their techniques and the factors they considered, and found the following:

Table 1. Literature Survey Analysis

Sr.No.	Paper	Technology	Summary
1	Fake Profile Identification using Machine Learning Algorithms. 2021	SVM algorithm, Random Forest Tree, Decision tree, Adaptive Boosting.	In this paper decision tree creates use of branch methodology to exemplify all within reach outcome of a call and supported sure conditions.
2	Detecting Fake account on social media using machine learning algorithms 2020	SVM-NN, NN, Random Forest	Here, SVM-NN, is used to provide effective detection, feature selection, and dimensionality reduction techniques for fake Twitter accounts and bots.
3	Detecting Clusters of Fake accounts in Online Social Networks 2019	Logistic Regression, Support Vector Machine	This study focuses on data mining, clustering, classification, spam detection, and fake profile identification.
4	Identify fake profile in online media 2021	Logistic Regression, support vector machine	The framework uses classification techniques to categorize profiles into fake or real categories
5	Fake Account Detection Using Machine Learning 2020	Gradient Boosting Machine. (GBM), XGBoost and AdaBoost	XGBoost, on the other hand, fared better with default values, achieving accuracy of up to 95%.

III. PROPOSED SYSTEM

There are many fake profiles online these days which generate false data and propaganda and also privacy violation, online bullying and trolling. So, our main aim is to eliminate all those fake accounts using system which can recognize if the account is fake or not before any activity of the user in the network or at the time of registration. In this system, we aim to develop a method by which we can detect fake accounts on twitter using machine learning. Our system is capable of classifying difference between a real account and fake account and thus. The reason for developing such a system is to identify fake accounts and thus stop cyber threats like spamming, spread of fake news agenda and There are now a lot more appearing on social media sites in various ways. there are many problems in today's online social networks such as fake profiles, online identity theft, etc. Even big social media sites like Facebook, twitter, Instagram are not able to solve this problem of fake accounts or bots in an efficient way.

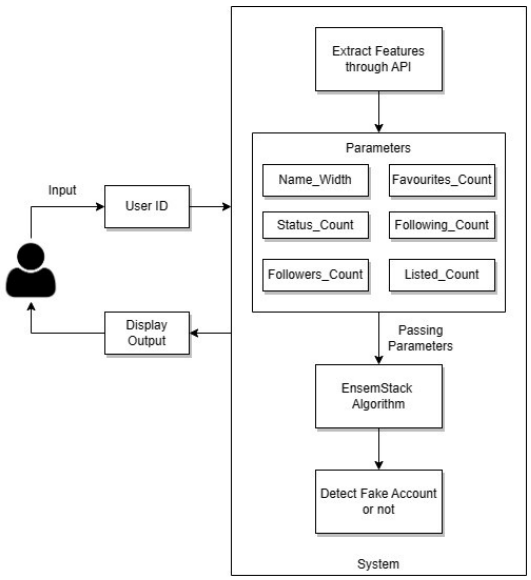

Fig. 1. Proposed SystemFlow

One of the reasons why world's richest man Elon Musk was about to cancel its Twitter deal was Twitter was unable to deliver proper reports about the fake accounts. In this project, we intend to create a mechanism to recognize bogus accounts so that people's social life is protected, and with the help of a fake account detection system, we can facilitate management by the websites of a large number of profiles that are not done Cannot be done manually.

The suggested system's architecture is depicted in Figure 1 together with a dataset of authentic and fraudulent Twitter accounts. After pre-processing the data by choosing pertinent features including name weight, statuses count, followers count, friends count, favorites count, and listed count, this dataset was used for training and assessing our machine learning model. We employed decision tree, neural network, and naive Bayes as our basic models and chose an ensemble technique called stacking that integrates them all. Instead of evaluating the performance of the stacking model using confusion matrix, these models were probably chosen based on how well they performed on the training data and how well they generalized to fresh data. In order to do this, predictions based on the test data are compared to the actual labels. To gauge the effectiveness of the model, metrics including accuracy, precision, recall, and F1-score are employed. Then, using the Twitter API, we created a web application that inputs user-specific Twitter account attributes into a pre-trained model to determine whether the account is legitimate or not. The website then shows the projected classification.

IV. SYSYTEM DESIGN

The first stage in the detection procedure is choosing the profile that has to be assessed. A feature is an attribute that affects or helps solve a problem, and selecting the key features for the model is referred to as feature selection. To achieve better results, a well-prepared and high-quality input dataset is required. To hone our model and aid in its learning, we gather a vast amount of data. In figure no.2 represents collection of datasets about 1482 rows 32 column and 1339 rows 32 column of real users and fake users respectively. After preprocessing the data, the dataset was 2819 rows and 6 columns. The dataset typically comprises of noisy data, useless data, and some usable data in small amounts. Additionally, the enormous volume of data slows down the model's training process, and with noise and irrelevant input, the model might not predict correctly or perform well. Therefore, it is imperative to eliminate these noises and less-important data from the dataset. To do this, feature selection techniques are employed. Making the optimal feature choices enhances the model's performance.

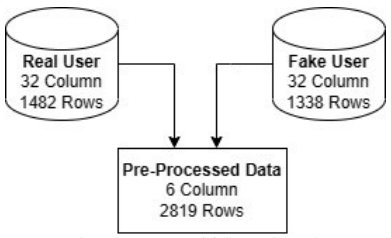

Fig. 2. Dataset of fake and Real User

After selecting the profile, selecting the required products and features, the next step involves dividing the data into two groups as mentioned in figure no.3. First, fit the model using a subset of the training dataset. The sample was not examined in the second subset. Instead, it takes input points from the dataset and compares its predictions with expected results. This is the benchmark dataset after the first dataset.

• Training Dataset: used to tune the machine learning model.

• Test Dataset: Used to evaluate tuned machine learning model.

Estimating how well the machine learning model performs on fresh data is the goal. Confusion Matrix will be used after that. A confusion matrix is a table that lets you see how well a classification model is performing. The data in it can also be used to compute metrics that will allow you to assess the model's usefulness.

While columns show the actual classes found in the data, rows show anticipated categories. Performance metrics for classification models can be computed using confusion matrices. The most used performance measurements are accuracy, precision, recall, and F1 score.

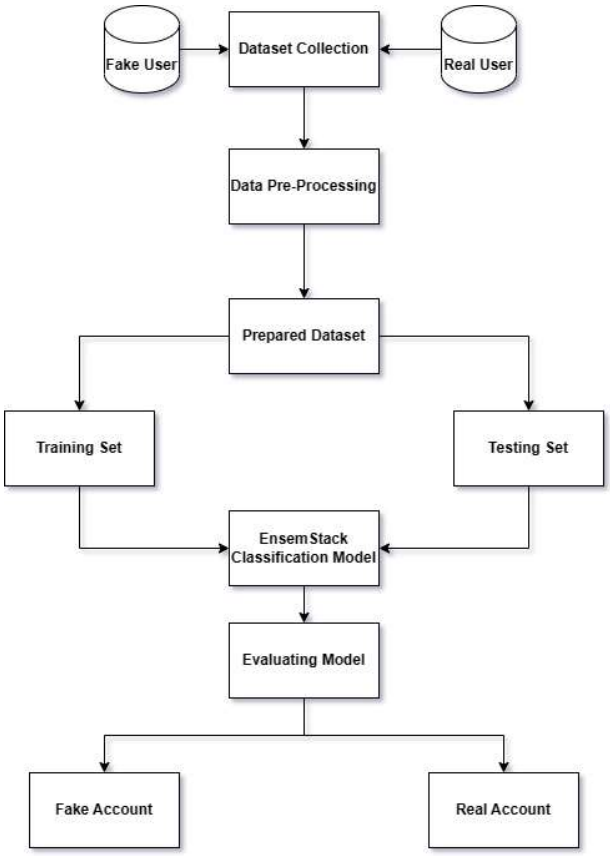

Fig. 3. Flowchart for design of system

V. Implementation Details

We will now use a Flask application that uses a machine learning model to estimate the probability of a Twitter user being identified. The main function of the application is my report, which uses input from a form on the web page, pulls data such as Name, Width, Status Count, Followers Count, Favorites Count, Following Count, Listed Count from Twitter's API and changes this information. It is processed as input to machine learning models and then predictions are made based on user data.
We import libraries useful for resizing or cleaning up our data, including NumPy, Panda, and Scikit-Learn. We also load Sklearn preprocessing to clean our data. Any machine learning model uses data preprocessing to produce better results, Although cleaning up data from raw data is done through data processing. We now manually select features for data preprocessing in the next section with better results. Our featured features are:

- name wt
- statuses count
- followers count
- friends count
- favorites count
- listed count

Process of constructing a stack ensemble model using three basic classifiers (decision tree, naive Bayes, and neural network) and a meta-classifier (single-layer neural network). Evaluate the ensemble model using the validation set and print the accuracy score.
Here is a breakdown of the code:

Step 1:
Import the necessary libraries.

Step 2:
load the Twitter dataset and assign it to variable data.

Step 3:
Extract all features except target variable labels and store them in a list called features.

Step 4:
Create training and validation sets from the dataset.

Step 5:
Define three base classifiers and their hyperparameters.

Step 6:
Defines a meta-classifier with a single hidden layer.

Step 7:
predict the target variable on the validation set and store the prediction.

Use the confusion matrix to improve the classification algorithm's performance. You may get a better understanding of where your classification model is accurate and what kind of inaccuracy there is by calculating the confusion matrix. The following are the calculating formulas:

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN}$$

$$\text{Precision} = \frac{TP}{TP+FP}$$

$$\text{Recall} = \frac{TP}{TP+FN}$$

$$\text{F-measure} = \frac{2*\text{Precision}*\text{Recall}}{\text{Precision}+\text{Recall}}$$

$$\text{Error Rate} = \frac{FP+FN}{TP+TN+FP+FN}$$

Table 2. Comparative analysis of different Models

Sr No.	Algorithm / Model Name	Accuracy
1.	Neural Network (NN)	82.29%
2.	Naive Bayes (NB)	89.18 %
3.	Decision Tree (DT)	93.15%
4.	Existing system[10]	90.41%
5.	Existing system[11]	96.28%
6.	Ensemble learning (Proposed Model)	98.93%

The above table represents the complete accuracies of all the tested and worked models. The basic classification algorithm like NN, NB, DT and Ensemstack have an accuracy of 82.29%, 89.18% ,93.15% and 98.93% respectively. We also referred other Reference Paper where the accuracy was 90.41% and 96.28%. Our proposed model, EnsemStack Classification Algorithm has an accuracy of 98.93% which is obtained among all the models. We'll now put into action a Flask application that uses a machine learning model to forecast how likely it is that a Twitter user will be verified. It uses Twitter's API to gather data after receiving input from a web form depends on user data such name length, number of status updates, followers, likes, follows, and lists, and analyses the data to feed machine learning models and show predictions on websites.

VI. Experiments and Results

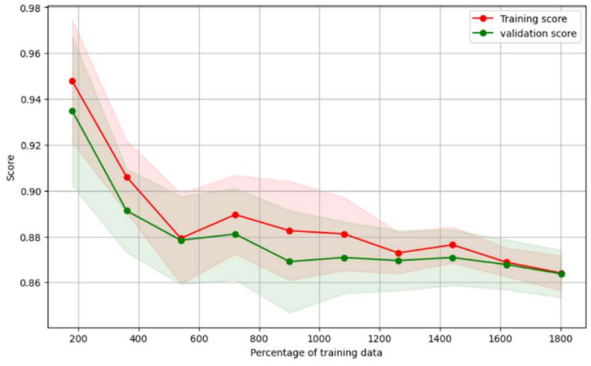

Fig. 4. Accuracy Graph of Naive Bayes Algorithm

Figure 4 shows how the neural network algorithm's score increased when more training data was fed into the model.

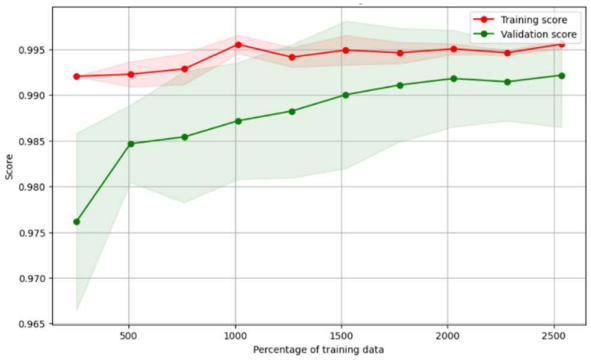

Fig. 5. Accuracy Graph of Decision Tree Algorithm

Figure 5 shows how the Decision Tree Algorithm's score increased when more training data was introduced into the model.

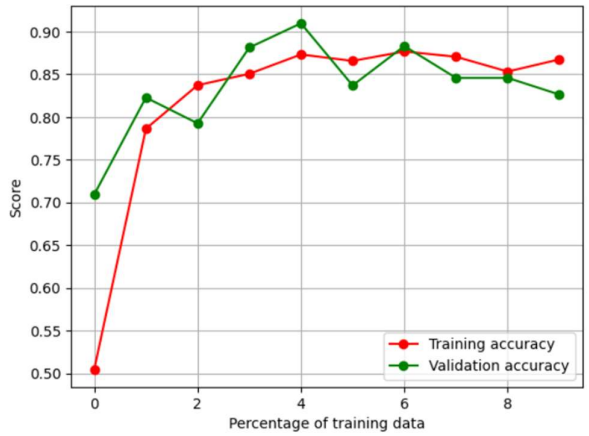

Fig. 6. Accuracy Graph of Neural Network Algorithm

Figure 6 shows how the Naive Bayes Algorithm's score increased when more training data was given into the model.

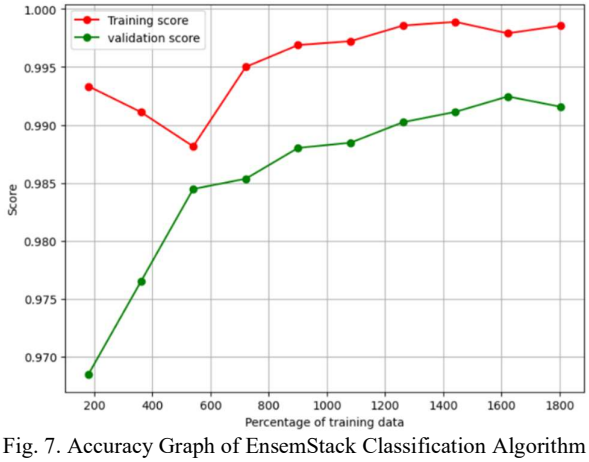

Fig. 7. Accuracy Graph of EnsemStack Classification Algorithm

Figure 7 shows how the Ensemble Stack Classification Algorithm's score increased when more training data was introduced into the model.

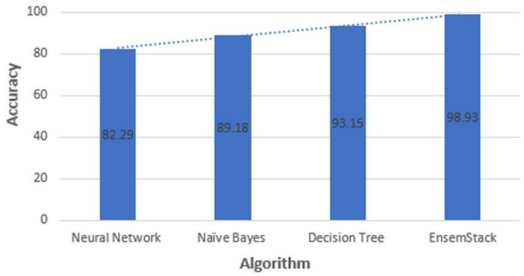

Fig. 8. Comparative analysis of different Models

Figure 8 compares the EnsemStack Classification Algorithm to all other classification models in terms of accuracy. Comparing the suggested system to prior existing models, it is more accurate. Both the data and the real-time application are fully supported. So, it will be capable of satisfying the requirements of the questions from the actual world.

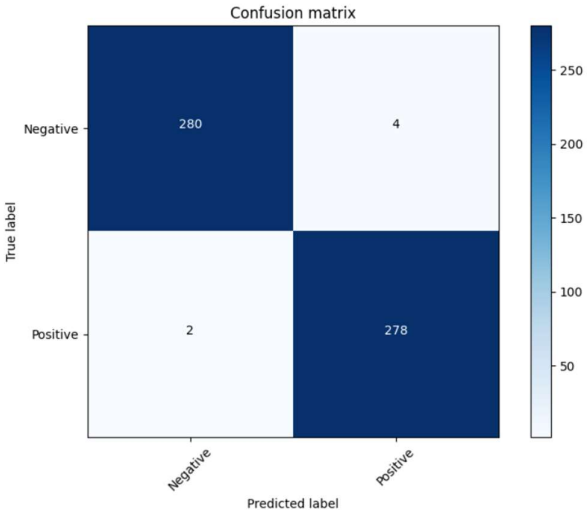

Fig. 9. Confusion Matrix for Stacking Classifier

By comparing the predicted labels to the actual labels, the confusion matrix in figure 9 highlights how well a classification model performed.

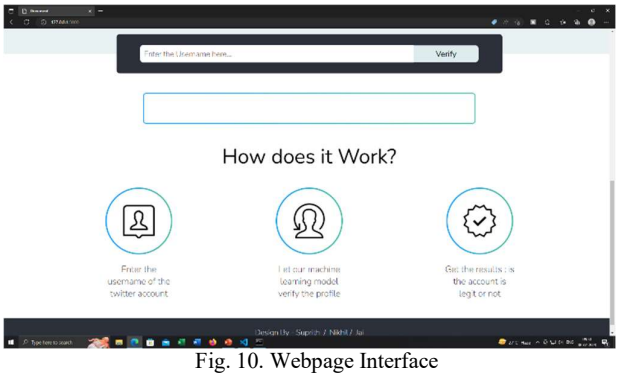

Fig. 10. Webpage Interface

The figure 10 represents we have created a easy to use web interface for our users by which they can check weather an account is fake or real by entering its it's user ID in search bar

Fig. 11. Real User

The figure 11 represents If the user profile is real then we will notify the user by saying it's an authentic profile and also providing the reasons for this particular classification.

Fig. 12. Fake User

The figure 12 represents If the user profile is fake then we will notify the user by saying it's a Fake profile and also providing the reasons for this particular classification.

VII. CONCLUSION

In conclusion, the fake Twitter account detection system is a powerful tool that can help combat the growing problem of fraudulent and misleading accounts on the platform. By leveraging machine learning technologies. The technology has the ability to evaluate massive amounts of data and spot irregularities and behaviors that might point to fabricated accounts. This can help Twitter users and moderators to take action against these accounts and to promote a more trustworthy and reliable online environment. However, it is important to note that the fake Twitter account detection system is not perfect and occasionally flag legitimate accounts as fake. As a result, it's critical to use the system in conjunction with human judgement and to continuously update and enhance the system's algorithms. Overall, the false Twitter account identification system is a useful instrument in the battle against online fraud and misinformation, and preserving the integrity and dependability of the Twitter network will depend on its continuing development and improvement.

REFERENCES

[1] Estee Van Der Walt and Jan Eloff,"Using Machine Learning to Detect Fake Identities:Bots vs Humans"IEEE Trans. Emerg.TopicsComput. Intell., vol. 1, no. 1, pp. 61–71 March 2018.

[2] Yel Chakraborty, Siddhartha Bhattacharyya, Rajib Bag, "A Survey of Sentiment Analysis from Social Media Data", IEEE Transactions on Computational Social Systems, Volume:7, Issue: 2, (2020)

[3] Muhammad Adil, Rahim Khan, M. Ahmad NawazUl Ghani, "Preventive Techniques of Phishing Attacks in Networks", 2020 3rd International Conference on Advancements in Computational Sciences (ICACS),(2020)

[4] Fatih Cagatay Akyon, M. Esat Kalfaoglu,"Fake Account Detection Using Machine Learning",2019 Innovations in Intelligent Systems and Applications Conference (ASYU),(2019)

[5] Ranojoy Barua, Rajdeep Maity, Dipankar Minj, Tarang Barua, Ashish Kumar Layek, "FNAD: An Application for Fake News Article Detection using Machine Learning Techniques", 2019IEEE Bombay Section Signature Conference (IBSSC), (2019)

[6] Estee Van Der Walt, Jan Eloff, "Using Machine Learning to Detect Fake Identities ",IEEE ´ Access, Volume: 6, (2018)

[7] SarahKhaled,Neamat El-Tazi and Hoda M. O. Mokhtar"Detecting Fake Accounts on Social Media" IEEE International Conference on Big Data.., vol.6 pp 101-110 ,2018.

[8] Gupta, Aditi, and Rishabh Kaushal."Towards detecting fake user accounts in facebook." 2017 ISEA Asia Security and Privacy (ISEASP). IEEE, 2017.

[9] AmedTorky, Ali Meligy and Hani Ibrahim"Recognizing Fake Identities In Online Social Networks Based on a FiniteAutomaton Approach"International Journal of Computer Applications, 2016.

[10] Buket Erúahin1 , Özlem Aktaú1 , Deniz KÕlÕnç2 , Ceyhun Akyol2 1 Computer Engineering Department, Dokuz Eylul University, Izmir, Turkey on Twitter Fake Account Detection

[11] Amna Kadhim Ali, Abdulhussein Mohsin Abdullah² Department of Computer Science, College of Computer Science and Information Technology, University of Basrah, Basrah, Iraq on Fake accounts detection on social media using stack ensemble system

[12] A. S. Chamria, A. D. Mane, P. V. Dambal and S. Bharne, "Detecting Fake Profile in Online Social Networks using EnsemStack Classification Algorithm," 2022 6th International Conference On Computing, Communication, Control And Automation (ICCUBEA, Pune, India, 2022, pp. 1-6, doi: 10.1109/ICCUBEA54992.2022.10010723.