IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

Fake Account Detection in Social Media

Khushboo Saraswat[1] and Nirupma Tiwari[2]

[1]Department of CSE, SRCEM, RGPV University Bhopal India

[2]Department of CSE, SRCEM, RGPV University Bhopal India

Abstract

Web-based media destinations are utilized consistently in these days and age, and have become a fundamental piece of our lives. It is one of the fundamental methods for correspondence, and has become an instrument for the two spammers and tricksters. Such web-based media stages have changed definitely how we carry on with our public activity. Making new friends, staying in touch with them, and keeping up with their activities have become much easier. Yet, numerous issues, like phony accounts and online pantomime, have additionally developed with the fast development of web-based media. Lately, web-based media stages, for example, Instagram, Twitter, and Facebook have bit by bit become significant approaches to scatter data. “One of these social stages that have pulled in more consideration in past years is Instagram. Instagram has generally utilized for sharing photographs and recordings and is beneficial for big names, organizations, and individuals with an extensive number of adherents. Meanwhile, this high benefit made this stage inclined to be the possible spot to be utilized for vindictive exercises. One of the fundamental malevolent exercises in the Instagram stage is Fake account.” The proprietors of Fake Account extricate the individual data about others and spread the produced information on interpersonal organizations. This Paper mainly focuses on Fake Account detection, using classification techniques (Random Forest, K-Nearest Neighbor, Neural Network with (Sigmoid Activation Function)) from machine learning. Our proposed model is capable of identifying the fake account with the possible minimum set of attributes.

Kaggle Instagram dataset has been used to examine the performance of implemented classifiers.

Keywords: Social Media, Fake accounts, Machine learning, classification, Instagram

I. INTRODUCTION

Social media platforms such as Instagram, Twitter, and Facebook have grown in popularity in recent years to disperse and share data [3]. These administrations give quick and reasonable correspondence and other kinds of instruments that cause their clients to have the option to straightforwardly share furthermore, distribute their media substance like pictures, recordings, also, sounds over the web [4]. Thus, other than the huge number of clients on friendly stages, these highlights and instruments have intrigued numerous digital hoodlums in utilizing them to perform their noxious exercises via web-based media stages adequately. Dissimilar to before, numerous assaults with a restricted or little impact would now be able to have a significant effect by utilizing on the web social stages [5]. Nonetheless, the impact of social media on people's lives is enormous, and many people utilise it to form more broad connections. [6].

Quite possibly the most famous web-based media is Instagram [7]. Instagram is a popular long-range informal communication tool designed for posting pictures and videos on the internet. It's similar to most other forms of social media in that those who make a record have an account and source of news, and can use that to share photos and recordings. Several major people and organizations have made their mark on Instagram in recent years; they use it to grow their business and fan base. [8]. Moreover, a significant number of them what's more, other acclaimed clients use it as a stage for publicizing.When someone's following exceeds a hundred thousand or millions, it's natural to draw on that person's track record as a part of team. Somewhat recently, numerous VIPs and customary individuals who came to a significant number of supporters on Instagram have utilized their records as a spot for publicizing. Individuals additionally attempt to build the number of their devotees for different reasons, for example, accomplishing more acclaim, being reliable, and being powerful.

As of May 2019, there are over a billion clients enrolled on Instagram. In the new years, Instagram has been discovered to utilize outsider applications, called bots [1]. While these can unquestionably mimic a client and stain their standing prompting 'fraud', there has additionally been more noteworthy occurrences of noxious methods of advancing the brand picture of an organization known as "influencer showcasing". Nowadays various organizations are utilizing web-based media to notice to their clients' necessities which has prompted one more negligence called Angler phishing [2].

Fake accounts are a pernicious danger to client security, as they can be abused to poach classified or then again close to home data, when worked by digital crooks imitating someone else. Thus, it is basic to distinguish such fake profiles before they represent a danger to client security[9].

From the start glimpse, the prevalence of a record is estimated by certain measurements, for example, devotee check or shared substance like the quantity of preferences, remarks or perspectives. Web-based media [11] are extraordinary stages for our lives, however there are various issues which should be tended to. Issues identified with web-based media, like classification, online maltreatment, abuse and harassing, and so forth are most normally utilized by Fake account that seem to have been created for the benefit of associations or people , which can harm notoriety and decrease the quantity of preferences and devotees of people. Fake account creation, on the other hand, is expected to do more harm than any other sort of digital malfeasance.

There are a few purposes behind making fake account via web-based media [10] presents a few reasons.

A few reasons why individuals make fake account are:

- Social Designing
- Online pantomime
- Advertising and Crusading
- Privacy Interruption and so forth

By and large, all web-based media spammers are lawful clients. It is thusly a test to remember them, as well as remembering them from lawful clients. In addition, fraudsters can in any case utilize modest robotized approaches. Obtaining legitimacy while maintaining faith and making it difficult for the vast majority of web-based media users to see. Online media counterfeit record ID is an order issue in which lawful clients are very much perceived from counterfeit client based on their relating highlights. Personality is a trait that is associated with an individual, aside from that person.

II. RELATED WORK

Today, web-based media is growing incredibly quick; these administrations are basic for some individuals in the public eye, particularly for promoting efforts and famous people and lawmakers who endeavor to advance themselves utilizing supporters and fans on web-based media [13]. Consequently, counterfeit records made for individuals and associations can be unsafe and harm to these individuals and organizations' notorieties lastly prompted the diminishing number of their genuine preferences and devotees. Also, a wide range of phony profiles adversely affect the benefits of online media for promoting and organizations in publicizing [14].These false profiles can be a conduit for online bullying; legitimate clients are also aware of safety in the online environment because of these false profiles [15]. As a result, many academics have investigated the subject of detecting malicious activities and spammers in the media using AI methodologies in recent years. In any case, there are a predetermined number of examination articles identifying with distinguishing counterfeit records or phony supporters. Inside this section, we discuss the two spammers and false account schemes that have recently revealed.

Yasyn Elyusufi et al.[2020]Using little profile data, this research developed a method for detecting a fraudulent profile on a social media site. For datasets that included both fraudulent and legal users, the proposed model was trained independently using the supervised learning technique. The ensemble classifier was employed to improve the accuracy of the predictions. In this study, three supervised machine learning methods are applied. To distinguish between fake and authentic profiles, Random Forest, Decision Tree, and Naive Bayes are utilised. With a precision score of 99.64 %, the Random Forest algorithm outperforms the other Algorithm.[11]

Fatih Cagatay Akyon , M. Esat Kalfaoglu [2019] As a dual grouping issue, this study discusses the identification of spurious and automated records that lead to a phony Instagram joint effort. This paper presented an anticipated work for false and robotized account recognition, as well as an expense sensitive element extraction approach based on a genetic computation for selecting the ideal characteristics for computerized account characterization. To distinguish phony and computerized accounts a few AI calculations like Guileless Bayes, strategic relapse, support Vector machines and neural organizations are utilized. The SVM and neural organizations acquired the best F1 score.SVM got 86 % and the neural organization got the most noteworthy F1 score with 95 %. [1]

BalaAnand et al. [2019] fostered another framework to identify counterfeit clients on the Twitter stage utilizing a diagram based semi-directed learning calculation (EGSLA) what's more, break down and assembling social and client created content (UGC) data. The model previously gathered clients' data, examined them to remove helpful highlights, and afterward performed grouping on these highlights and decided. The test results show that the EGSLA calculation accomplished superior and was more advantageous than other calculations, for example, Decision tree, KNN, SVM, and game hypothesis based techniques as far as grouping precision. [16]

Farhan Nurdiatama Pakaya et al. [2019] All through this work, an order model has been created utilizing just record tweets that separate real records from counterfeit records. Four unique calculations are utilized in this paper like Logistic Regression, ADA Boost, XG Boost, and Random Forest. Model assessment of this exploration uncovers that XG Boost with tf-idf highlights is the best format for the parallel order plot and the best model for multiclass grouping is world2vec highlights. This work has figured out how to accomplish a greatest precision of 95.5 %. [19]

Zulfikar Alom et al.[2018] Proposed another and all the more remarkable assortment of highlights to distinguish twitter spammers. It showed both diagram based and tweet-based attributes and applied them to seven distinctive AI calculations, like K-NN, Decision Tree, Naive Bayes, Random Forest, Logistic Regression, Support Vector and Extreme Gradient Boosting (XG-Boost). The investigation uncovers that the Random Forest calculation accomplishes a superior exhibition contrasted with different calculations with an exactness of 91%. [20]

Ahmed El Azab et al.[2016] A characterization technique for distinguishing counterfeit records on twitter is available in this paper. The investigation distinguishes a limited arrangement of key factors that impact the ID of a phony record on twitter and afterward concludes which are applied utilizing diverse characterization methods. In this paper, the arrangement technique utilized are Random Forest, Decision Tree, Naive Bayes, neural network, Support vector Machine. [18]

G. Supraja et al. [2015] This paper presents the pattern detection approach to detect the fake accounts. In this paper, the crawler is used to collect the twitter dataset. The pattern matching algorithm is used on the screen name and updates time of tweets to detect a group of fake accounts. The time to create the profile is analyzed for detection. The time taken by the fake user is different than the real user. The advantage of this approach is that it is a fast approach to detect the fake accounts. The disadvantage of this approach is that it detects only fewer numbers of fake accounts. [17]

G.Magno et al. [2010] This paper presents the issue to recognizing spammers on twitter. In this dataset of twitter is gathered and named the pre-characterized spammer and non-spammer users. Then credits are recognized dependent on the social conduct of the client. In this paper directed AI strategy SVM is utilized to find the spammers. Outspread Premise Capacity (RBF) part of Nonlinear SVM is utilized orders extremely complex data. Based on the ten credits this procedure separates the spammer and non-spammers. In this 70% of spammers and 96% of non-spammers are properly perceived. The methodology of this paper is likewise ready to distinguish spam as an option of spammers. The exactness of recognizing spam is 87.2%. [12]

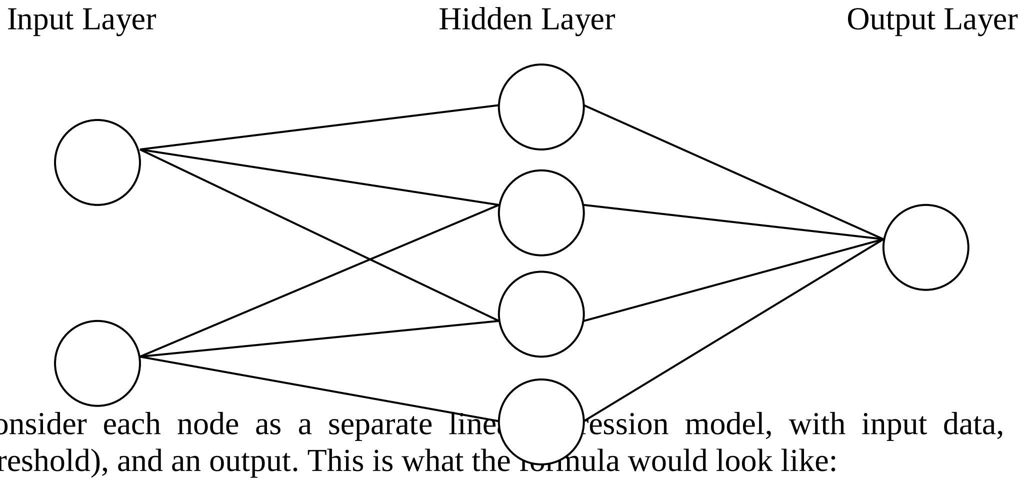
III. ACTION CLASSIFICATION METHODS

A. Machine learning approaches: Machine learning approaches used in our work are Random Forest, K-Nearest Neighbor, and Neural Network with (Sigmoid Activation Function)

- Random Forest: Random Forest is a well-known machine learning algorithm that uses the supervised learning method. In machine learning, it can be utilized for both classification and regression issues. It is based on ensemble learning, which is a method of integrating several classifiers to solve complex issues and achieve the model's performance.
  As the name recommends, "Random Forest is a classifier that contains various decision trees on different subsets of the given dataset and takes the normal to improve the prescient exactness of that dataset." Rather than depending on one decision tree, the Random Forest takes the forecast from each tree and dependent on the lion's share votes of expectations, and it predicts the last yield.

- K-Nearest Neighbor (KNN): The K-Nearest Neighbor algorithm is based on the Supervised Learning technique and is one of the most basic Machine Learning algorithms. The K-NN method assumes that the new case/data and existing cases are similar and places the new case in the category that is most similar to the existing categories. The K-NN method stores all available data and classifies a new data point based on its similarity to the existing data. This means that new data can be quickly sorted into a well-defined category using the K-NN method.
  The K-NN approach can be used for both regression and classification, but it is more commonly utilized for classification tasks.

- Neural Network: Neural Network is also known as Artificial neural networks (ANNs) and simulated neural networks (SNNs) are a subset of machine learning that are at the heart of deep learning methods. Their name and structure are derived from the human brain, and they resemble the way biological neurons communicate with one another.
  A node layer contains an input layer, one or more hidden layers, and an output layer in artificial neural networks (ANNs). Each node, or artificial neuron, is connected to the others and has a weight and threshold linked with it. If a node's output exceeds a certain threshold value, the node is activated, and data is sent to the next tier of the network. Otherwise, no data is sent on to the network's next tier.



Consider each node as a separate linear regression model, with input data, weights, a bias (or threshold), and an output. This is what the formula would look like:

$$\sum wixi + bias = w1x1 + w2x2 + w3x3 + bias$$

$$Output = f(x) = 1 \text{ if } \sum w1x1 + b >= 0;\ 0 \text{ if } \sum w1x1 + b < 0$$

Activation function: Activation function decides, whether a neuron should be activated or not by calculating weighted sum and further adding bias with it. The purpose of the activation function is to introduce non-linearity into the output of a neuron.

VARIANTS OF ACTIVATION FUNCTION:-

1). Linear Function:-

Equation: Linear functions have an equation that is comparable to that of a straight line, i.e. y = ax
Range: -inf to +inf
Uses: The linear activation function is only used once, in the output layer.
Issues: If we differentiate a linear function to introduce non-linearity, the outcome will no longer be dependent on the input "x" and the function will become constant, hence our algorithm will not exhibit any novel behavior.

2). Sigmoid Function: - It's a function that's graphed as a 'S' shape.

Equation: A = 1/(1 + e-x)
Nature: Non-linear in nature. The X values range from -2 to 2, but the Y values are extremely steep. This indicates that slight changes in x will result in huge changes in Y's value.
Value Range: 0 to 1
Uses: Typically employed in the output layer of a binary classification, where the result is either 0 or 1. Because the sigmoid function's value is only between 0 and 1, the result can be easily anticipated to be 1 if the value is greater than 0.5, and 0 otherwise.

3). Tanh Function:-

Equation :- f(x) = tanh(x) = 2/(1 + e-2x) - 1

OR

tanh(x) = 2 * sigmoid(2x) – 1

Value Range:- -1 to +1

Nature:- non-linear
Uses: Usually employed in hidden layers of a neural network since its values range from -1 to 1, causing the hidden layer's mean to be 0 or very close to it, which aids in data centering by bringing the mean close to 0.

4). RELU: - Stands for Rectified linear unit.

Equation: - A(x) = max(0,x).
Value Range: - [0, inf)
Nature: - non-linear,
Uses:- Because it includes fewer mathematical calculations, ReLu is less computationally expensive than tanh and sigmoid. Only a few neurons are active at a time, making the network sparse and efficient for computation.

5). Softmax Function:- The softmax function is a type of sigmoid function that comes in handy when dealing with classification issues.

Nature :- non-linear
Uses: This is commonly used when dealing with various classes. The softmax function would divide by the sum of the outputs and squeeze the outputs for each class between 0 and 1.

IV. PROPOSED MODEL

Algorithm:

Step1. Select validate Dataset
The first step of detecting the fake account on Instagram is collect the data set of Instagram .

Step2. Pre-process the dataset by detecting and managing the missing values.
Pre-processing refers to the changes we make to our data before feeding it to the algorithm. Data preprocessing is a technique for transforming unclean data into clean data. To put it another way, whenever data is acquired from numerous sources, it is in raw format, making analysis impossible. There was no missing data in the dataset. Missing values in a dataset can occur owing to a variety of factors, the most common of which are real-world issues. Missing values can be managed by deleting or imputation. The existence of missing values limits the amount of data that can be analyzed, lowering the study's statistical power and, as a result, the study's conclusions' dependability.

Step3. Divide pre-processed data into two parts (70% training data & 30% testing data).
To assess how effectively our machine learning model works, we must divide a dataset into train and test sets. The train set is used to fit the model, and the train set's statistics are well-known. The test data set is the second collection, and it is only utilised to make predictions.

Step4. Apply various machine learning approaches on data.
After preprocessing the classification approach is used. For data classification, (Random Forest (RF), K-Nearest Neighbor (KNN), Neural Network with (Sigmoid Activation Function)) are applied. These strategies are the most successful, and they consistently outperform other algorithms in terms of accuracy.

Step5. Select the most accurate classification model.

Classification techniques used in both the existing and proposed work, generate distinct outcomes. Strategies findings are compared to those of other techniques. The parameters used and the time required by the existing and proposed techniques are compared. The technique that produced results with higher accuracy is considered as the best classification model.

Step6. End

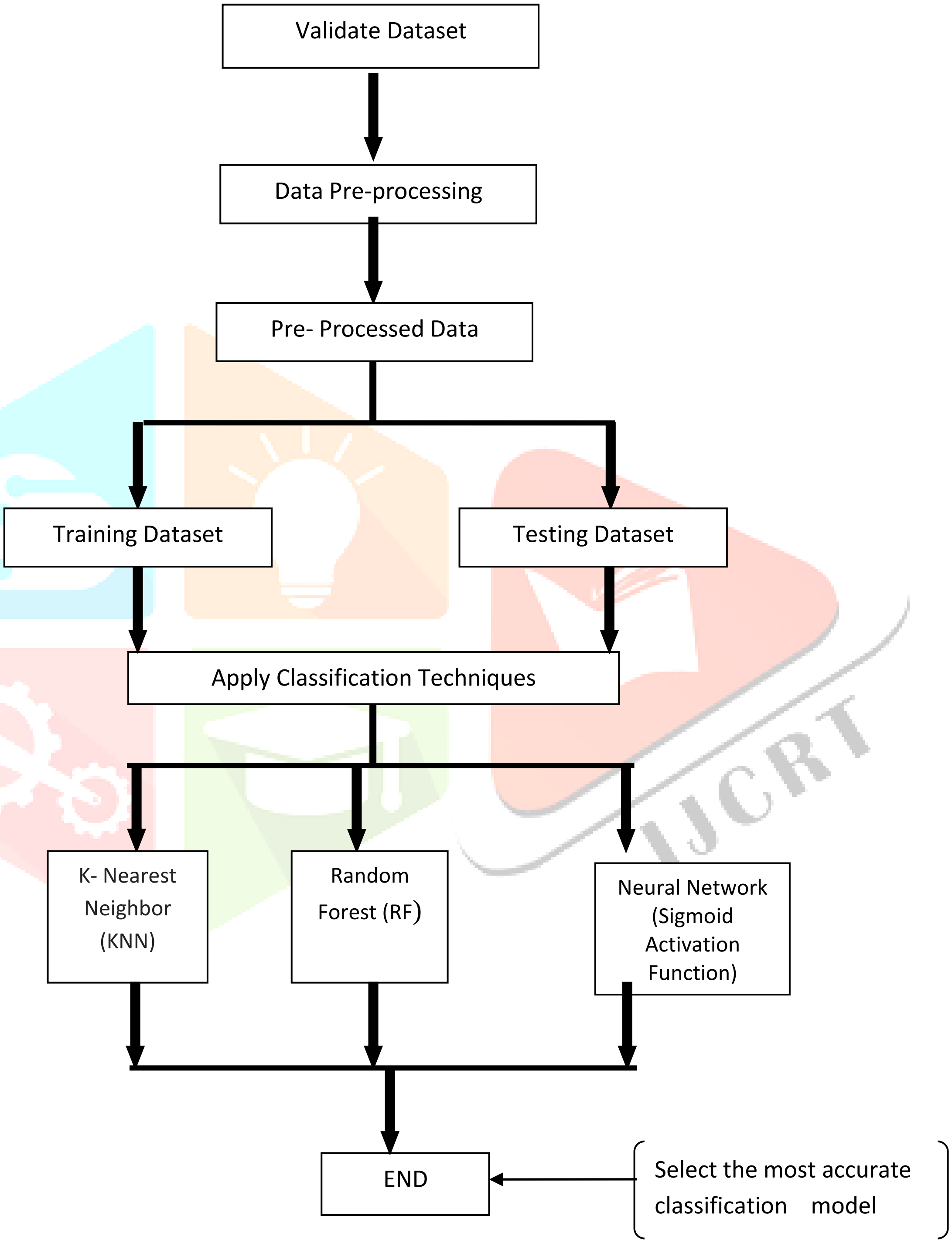

Validate Dataset
Data Pre-processing
Pre- Processed Data
Training Dataset
Testing Dataset
Apply Classification Techniques
K- Nearest Neighbor (KNN)
Random Forest (RF)
Neural Network (Sigmoid Activation Function)
END
Select the most accurate classification model

Figure2. The flowchart of Fake Account Detection

B. Dataset Description

The dataset has been taken fromhttps://www.kaggle.com/datasets/free4ever1/instagram-fake-spammer-genuine-accounts. It comprises of two CSV documents train.csv and test.csv .The reliant variable, which is if it is a phony record is clear cut and it takes two qualities 0 (not fake) and 1 (fake) profile. The conveyance of the preparation dataset is to such an extent that half is fake and the rest half is real. In dataset the table is to signify the boundaries that have been thought of (meant in section Profile include), their scope of qualities, every one of their mean qualities and what every one of the highlights denote.

C. Implementation: In this work, Fake account detection is based on various features using Python and applying various supervised multiclass classification machine learning algorithms.

V. RESULT AND DISCUSSION

The performance findings are contrasted with current methods that were evaluated on the datasets. Neural Network model is used for better prediction, after examining the existing model.

Confusion matrix :The performance of a classification model is measured using a N x N matrix, where N is the number of class labels. The matrix compares the actual target values to the predictions of the machine learning model. This gives us a clear picture about how well our classification method is performing and what kinds of errors it makes. In the confusion matrix

- The target variable can have either a positive or negative value.
- The actual values of the target variable are shown in the columns.
- The rows represent the predicted values for the target variable.

We'd use a 2 × 2 matrix with four values for a binary classification problem, as seen below:

True Positive (TP)

- The real and projected values are similar.
- The model predicted a positive result, and the outcome was positive.

True Negative (TN)

- The real and projected values are identical.
- The model predicted a negative result, and the result was negative.

False Positive (FP)

- The predicted value was inaccurate.
- The actual number was negative, while the model anticipated that it would be positive.

False Negative (FN)

- The value that was projected was inaccurate.
- The model anticipated a negative value, even though the actual result was positive.

Confusion matrix for all the methods used in this work are shown below:

(a) Random Forest (RF): Confusion matrix for this method is shown

	POSITIVE	**NEGATIVE**
TRUE	125	19
FALSE	39	106

Table1. Confusion matrix of RF

(b) K- Nearest Neighbor (KNN) : Confusion matrix for this method is shown

	POSITIVE	**NEGATIVE**
TRUE	129	15
FALSE	40	105

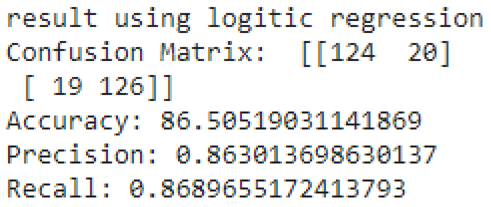
Table2. Confusion matrix of KNN

(c) Neural Network With (Sigmoid Activation Function) : Confusion matrix for this method is shown

	POSITIVE	**NEGATIVE**
TRUE	47	5
FALSE	12	52

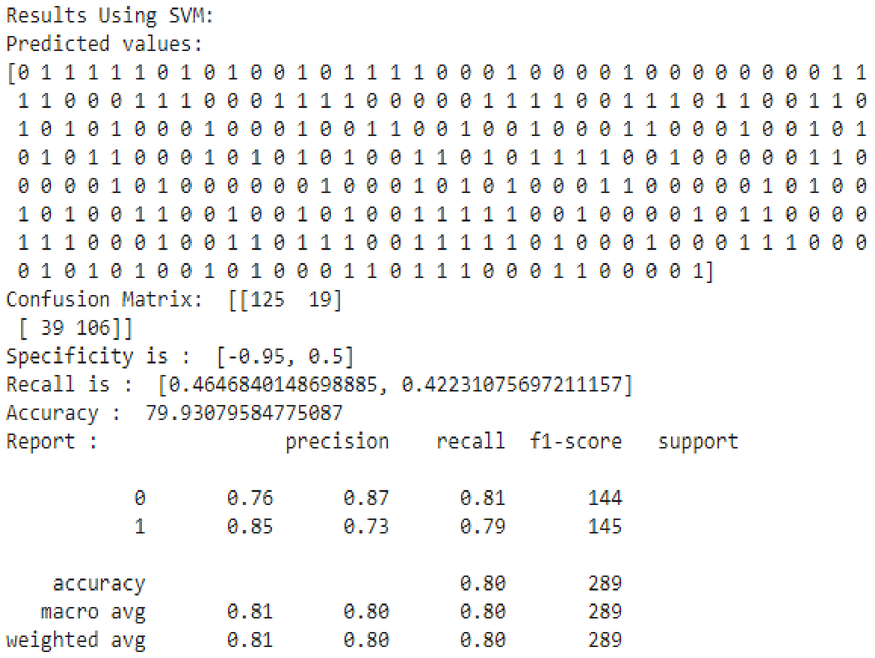
Table2. Confusion matrix of Neural Network with (Sigmoid Activation Function)

Accuracy of Existing work using Logistic Regression model:



```
result using logitic regression
Confusion Matrix:  [[124  20]
 [ 19 126]]
Accuracy: 86.50519031141869
Precision: 0.863013698630137
Recall: 0.8689655172413793
```

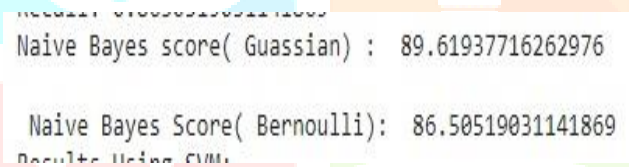
Accuracy of Existing work using Support Vector Machine model:



```
Results Using SVM:
Predicted values:
[0 1 1 1 1 1 0 1 0 1 0 0 1 0 1 1 1 1 0 0 0 1 0 0 0 0 1 0 0 0 0 0 0 0 0 1 1
 1 1 0 0 0 1 1 1 0 0 0 1 1 1 1 0 0 0 0 0 1 1 1 1 0 0 1 1 1 0 1 1 0 0 1 1 0
 1 0 1 0 1 0 0 0 1 0 0 0 1 0 0 1 1 0 0 1 0 0 1 0 0 0 1 1 0 0 0 1 0 0 1 0 1
 0 1 0 1 1 0 0 0 1 0 1 0 1 0 1 0 0 1 1 0 1 0 1 1 1 1 0 0 1 0 0 0 0 0 1 1 0
 0 0 0 0 1 0 1 0 0 0 0 0 0 1 0 0 0 1 0 1 0 1 0 0 0 1 1 0 0 0 0 0 1 0 1 0 0
 1 0 1 0 0 1 1 0 0 1 0 0 1 0 1 0 0 1 1 1 1 1 0 0 1 0 0 0 0 1 0 1 1 0 0 0 0
 1 1 1 0 0 0 1 0 0 1 1 0 1 1 1 0 0 1 1 1 1 1 0 1 0 0 0 1 0 0 0 1 1 1 0 0 0
 0 1 0 1 0 1 0 0 1 0 1 0 0 0 1 1 0 1 1 1 0 0 0 1 1 0 0 0 0 1]
Confusion Matrix:  [[125  19]
 [ 39 106]]
Specificity is :  [-0.95, 0.5]
Recall is :  [0.4646840148698885, 0.42231075697211157]
Accuracy :  79.93079584775087
Report :               precision    recall  f1-score   support

           0       0.76      0.87      0.81       144
           1       0.85      0.73      0.79       145

    accuracy                           0.80       289
   macro avg       0.81      0.80      0.80       289
weighted avg       0.81      0.80      0.80       289
```

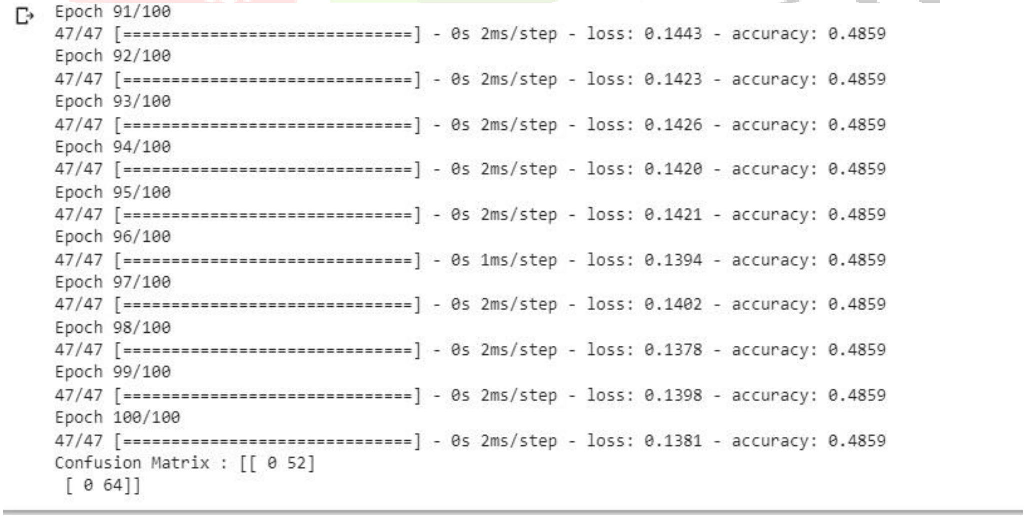
Accuracy of Existing work using Naive Bayes (Guassian) Model

Accuracy of Existing work using Naive Bayes (Bernoulli) Model



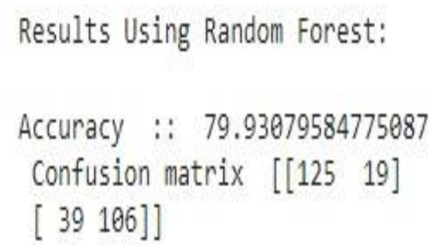
```
Naive Bayes score( Guassian) :  89.61937716262976

 Naive Bayes Score( Bernoulli):  86.50519031141869
```

Accuracy of Existing work using Neural Network with (Softmax Activation Function) Model



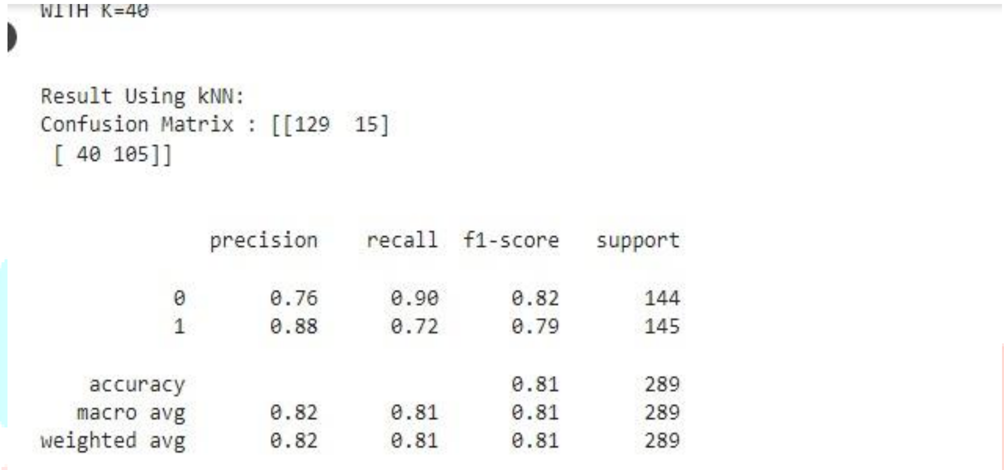
```
Epoch 91/100
47/47 [==============================] - 0s 2ms/step - loss: 0.1443 - accuracy: 0.4859
Epoch 92/100
47/47 [==============================] - 0s 2ms/step - loss: 0.1423 - accuracy: 0.4859
Epoch 93/100
47/47 [==============================] - 0s 2ms/step - loss: 0.1426 - accuracy: 0.4859
Epoch 94/100
47/47 [==============================] - 0s 2ms/step - loss: 0.1420 - accuracy: 0.4859
Epoch 95/100
47/47 [==============================] - 0s 2ms/step - loss: 0.1421 - accuracy: 0.4859
Epoch 96/100
47/47 [==============================] - 0s 1ms/step - loss: 0.1394 - accuracy: 0.4859
Epoch 97/100
47/47 [==============================] - 0s 2ms/step - loss: 0.1402 - accuracy: 0.4859
Epoch 98/100
47/47 [==============================] - 0s 2ms/step - loss: 0.1378 - accuracy: 0.4859
Epoch 99/100
47/47 [==============================] - 0s 2ms/step - loss: 0.1398 - accuracy: 0.4859
Epoch 100/100
47/47 [==============================] - 0s 2ms/step - loss: 0.1381 - accuracy: 0.4859
Confusion Matrix : [[ 0 52]
 [ 0 64]]
```

Accuracy of Proposed work using Random Forest (RF) model:



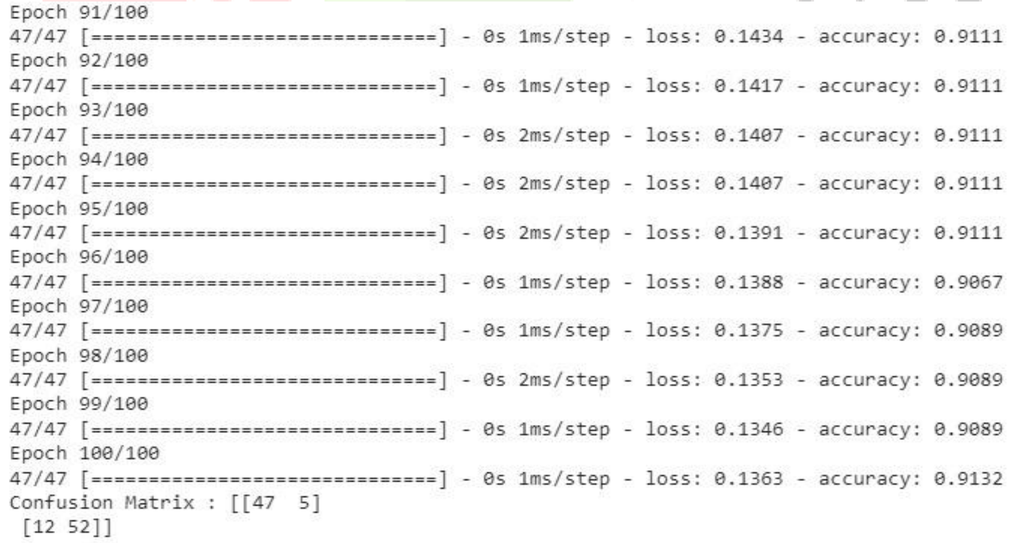
```
Results Using Random Forest:

Accuracy  ::  79.93079584775087
 Confusion matrix  [[125  19]
 [ 39 106]]
```

Accuracy of Proposed work using K- Nearest Neighbor (KNN) model:



```
WITH K=40

Result Using kNN:
Confusion Matrix : [[129  15]
 [ 40 105]]


              precision    recall  f1-score   support

           0       0.76      0.90      0.82       144
           1       0.88      0.72      0.79       145

    accuracy                           0.81       289
   macro avg       0.82      0.81      0.81       289
weighted avg       0.82      0.81      0.81       289
```

Accuracy of Proposed work using Neural Network with Sigmoid Activation Function



```
Epoch 91/100
47/47 [==============================] - 0s 1ms/step - loss: 0.1434 - accuracy: 0.9111
Epoch 92/100
47/47 [==============================] - 0s 1ms/step - loss: 0.1417 - accuracy: 0.9111
Epoch 93/100
47/47 [==============================] - 0s 2ms/step - loss: 0.1407 - accuracy: 0.9111
Epoch 94/100
47/47 [==============================] - 0s 2ms/step - loss: 0.1407 - accuracy: 0.9111
Epoch 95/100
47/47 [==============================] - 0s 2ms/step - loss: 0.1391 - accuracy: 0.9111
Epoch 96/100
47/47 [==============================] - 0s 1ms/step - loss: 0.1388 - accuracy: 0.9067
Epoch 97/100
47/47 [==============================] - 0s 1ms/step - loss: 0.1375 - accuracy: 0.9089
Epoch 98/100
47/47 [==============================] - 0s 2ms/step - loss: 0.1353 - accuracy: 0.9089
Epoch 99/100
47/47 [==============================] - 0s 1ms/step - loss: 0.1346 - accuracy: 0.9089
Epoch 100/100
47/47 [==============================] - 0s 1ms/step - loss: 0.1363 - accuracy: 0.9132
Confusion Matrix : [[47  5]
 [12 52]]
```

The table3 below shows the findings of the Existing work of the data mining techniques for performance assessment, namely the Support vector machine, Logistic Regression, Naive Bayes (Gaussian) , Naive Bayes (Bernoulli) , Neural Network (Softmax Activation Function) method.

Models	Existing work (%)
SVM (support vector machine)	79.93
LR(Logistic Regression)	86.50
Naive Bayes (Gaussian)	89.61
Naive Bayes (Bernoulli)	86.50
Neural Network (Softmax Activation Function)	48.59

Table3. Performance Findings of the existing work of tested models

The table4 below shows the findings of the proposed work of the data mining techniques for performance assessment, namely the Random Forest, K-Nearest Neighbor, and Neural Network with (Sigmoid Activation Function) method.

Models	Proposed work (%)
Random Forest (RF)	79.93
K- Nearest Neighbor (KNN)	81
Neural Network (Sigmoid Activation Function)	91.32

Table4. Performance Findings of the Proposed work of tested models

The table5 below shows the comparative findings of the data mining techniques used in Base Paper work and Proposed Work for performance assessment.

Base Paper Work Model		**Proposed Work Model**	
Models	**Accuracy (%)**	**Models**	**Accuracy (%)**
SVM (support vector machine)	79.93	Random Forest (RF)	79.93
LR(Logistic Regression)	86.50	K- Nearest Neighbor (KNN)	81
Naive Bayes (Gaussian)	**89.61**	Neural Network (Sigmoid Activation Function)	91.32
Naive Bayes (Bernoulli)	86.50	-	-
Neural Network (Softmax Activation Function)	48.59	-	-

Table5. Comparison of the existing and proposed performance of tested models

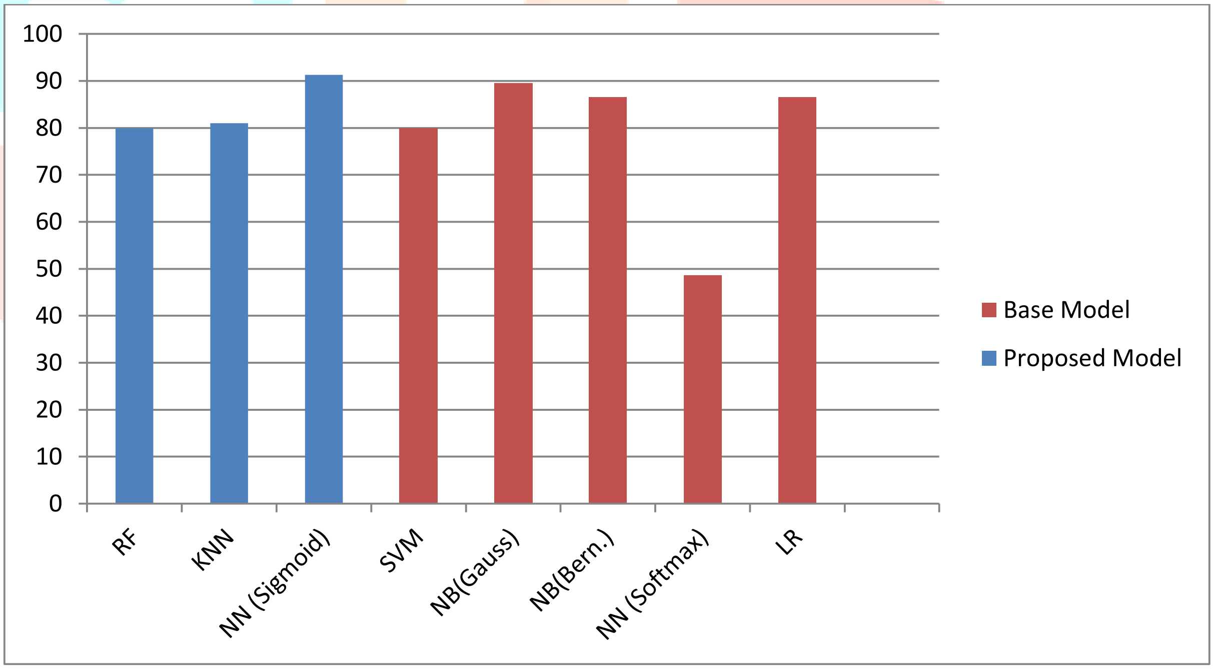


Figure3. Graph representation of proposed work and Existing Work (Accuracy)

VI. CONCLUSION

As shown in table4, the more precise technique tested in the paper is Neural Network (Sigmoid Activation Function) with highest accuracy score. We demonstrated the application of supervised classification machine learning algorithms for detecting fake accounts on social media in this study**.** While reviewing prior comparable studies on the identification of false profiles on social media sites, we discovered that not much has been done on Instagram in particular. As a result, we tailored our strategy to the situation. In this research, we presented a unique technique for identifying false user accounts on Instagram based on certain attributes utilising machine learning concepts. For this, we employed three models: Random Forest, K-Nearest Neighbor, and Neural Network with (Sigmoid Activation Function) with accuracy rates of 79.93 percent, 81 percent and 91.32 percent, respectively.

REFERENCES

1. F. C. Akyon and M. Esat Kalfaoglu, "Instagram Fake and Automated Account Detection," 2019 Innovations in Intelligent Systems and Applications Conference (ASYU), Izmir, Turkey, 2019, pp. 1-7, doi: 10.1109/ASYU48272.2019.8946437.
2. Ananya Dey , Hamsashree Reddy , Manjistha Dey and Niharika Sinha , Detection of Fake Accounts in Instagram using Machine Learning, Vol 11, No 5, October 2019.
3. Sheikhi, S. (2020). An efficient method for detection of fake accounts on the Instagram platform. Revue d'Intelligence Artificielle, Vol. 34, No. 4, pp. 429-436. https://doi.org/10.18280/ria.340407.
4. Boididou, C., Middleton, S.E., Jin, Z., Papadopoulos, S., Dang-Nguyen, D.T., Boato, G., Kompatsiaris, Y. (2018). Verifying information with multimedia content on twitter. Multimedia Tools and Applications, 77(12): 15545- 15571. https://doi.org/10.1007/s11042-017-5132-9.
5. Alqatawna, J., Madain, A., Al-Zoubi, A., Al-Sayyed, R. (2017). Online social networks security: Threats, attacks, and future directions. In Social Media Shaping ePublishing and Academia, pp. 121-132. https://doi.org/10.1007/978-3-319-55354-210.
6. 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
7. 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. https://doi.org/10.1016/j.jretconser.2019.03.012
8. 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
9. A. Narayanan, A. Garg, I. Arora, T. Sureka, M. Sridhar and H. B. Prasad, "IronSense: Towards the Identification of Fake User-Profiles on Twitter Using Machine Learning," 2018 Fourteenth International Conference on Information Processing (ICINPRO), 2018, pp. 1-7, doi: 10.1109/ICINPRO43533.2018.9096687.
10. Bharat Sampatrao Borkar, Dr. Rajesh Purohit (2019).Recognisation of fake profiles in social media. Department of Computer Science & Engineering School of Engineering & Technology, Suresh Gyan Vihar University, Jagatpura, Volume-9 Issue-2,2019.
11. Elyusufi, Yasyn & Elyusufi, Zakaria & M'hamed, Aït Kbir. (2020). Social Networks Fake Profiles Detection Using Machine Learning Algorithms. 10.1007/978-3-030-37629-1_3.
12. G. Magno and T. Rodrigues, “Detecting Spammers on Twitter,” CEAS, 2010
13. Blair, S.J., Bi, Y., Mulvenna, M.D. (2020). Aggregated topic models for increasing social media topic coherence. Applied Intelligence, 50(1): 138-156. https://doi.org/10.1007/s10489-019-01438-z
14. Jiang, X., Li, Q., Ma, Z., Dong, M., Wu, J., Guo, D. (2019). QuickSquad: A new single-machine graph computing framework for detecting fake accounts in large-scale social networks. Peer-to-Peer Networking and Applications, 12(5): 1385-1402. https://doi.org/10.1007/s12083-018-0697-2
15. Ramalingam, D., Chinnaiah, V. (2018). Fake profile detection techniques in large-scale online social networks: A comprehensive review. Computers & Electrical Engineering, 65: 165-177. https://doi.org/10.1016/j.compeleceng.2017.05.020

16. BalaAnand, M., Karthikeyan, N., Karthik, S., Varatharajan, R., Manogaran, G., Sivaparthipan, C.B. (2019). An enhanced graph-based semi-supervised learning algorithm to detect fake users on Twitter. The Journal of Supercomputing, 75(9): 6085-6105. https://doi.org/10.1007/s11227-019-02948-w
17. B. Hudson, J. Matthews, S. Gurajala, J. S. White, B. Hudson, and J. N. Matthews, “Fake Twitter accounts : Profile characteristics obtained using an activity-based 50 pattern detection approach Fake Twitter accounts : Profile characteristics obtained using an activity-based pattern detection approach,” ACM, no. August, 2015
18. Elazab, Ahmed & Mahmood, Mahmood & Hefny, Hesham. (2016). Fake Account Detection in Twitter Based on Minimum Weighted Feature set. International Journal of Computer, Electrical, Automation, Control and Information Engineering Vol:10, No:1, 2016.
19. F. N. Pakaya, M. O. Ibrohim and I. Budi, "Malicious Account Detection on Twitter Based on Tweet Account Features using Machine Learning," 2019 Fourth International Conference on Informatics and Computing (ICIC), Semarang, Indonesia, 2019, pp. 1-5, doi: 10.1109/ICIC47613.2019.8985840.
20. Z. Alom, B. Carminati and E. Ferrari, "Detecting Spam Accounts on Twitter," 2018 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM), Barcelona, 2018, pp. 1191-1198, doi: 10.1109/ASONAM.2018.8508495.