Artificial Intelligence Tool for Fake Account Detection from Online Social Networks

Dr. B. Subba Reddy[1], D. R. Amrutha Nayana[2], G. Sahaja[2], G. Shanmukha Priya[2]

[1]*Professor & HOD,* [2]*UG Student,* [1,2]*Department of Information Technology*

[1,2]*Malla Reddy Engineering College for Women (UGC-Autonomous), Maisammaguda, Secunderabad, Telangana, India*

Abstract

Online social networks such as Facebook or Twitter contains user's details, and some malicious users will hack social network database to steal or breach users' information. Usually, all fake user's main intention is to send friend request to normal users to hack their machine or to steal their data and never they will have many numbers of posts or have many following friends and their account age also will have a smaller number of years. By analysing this features Facebook will mark whether user profile is fake or genuine. Therefore, to protect the users' data, this project uses artificial neural networks (ANNs) to identify whether given social network account details are from genuine or fake users. ANN algorithm will be trained with all previous users fake and genuine account dataset then this ANN pretrained model will be applied on new test data to identify whether the given new account details are from genuine or fake users.

Keywords: online social media, fake account detection, artificial neural networks.

1. Introduction

Online social media is the place each person has an outlook then be able to keep connecting their relations, transfer their updates, join with the people having same likes. Online Social Networks makes use of front-end technologies, which permits permanency accounts in accordance with to know each other. Facebook and Twitter are developing along with humans to maintain consultation together with all others. The online accounts welcome people including identical hobbies collectively who makes users easier after performing current friends. Gaming and entertaining web sites which have extra followers unintentionally that means more fan base and supreme ratings [1]. Ratings drives online account holders to understand newer approaches not naturally or manually to compete more with their neighbours. By these analogies, the maximum famous candidate in an election commonly gets a greater number of votes. Happening of fake social media accounts and interests may be known [2]. Instance is fake online account being sold on-line at an online marketplace for minimum price, brought from collaborative working offerings. More often feasible to have Twitter fans and Facebook media likes in online. Fake user accounts may be created by humans or computers like bots, cyborgs. Cyborg is half bot and half human account. These accounts are usually opened by human, but their actions are made by bots. Another reason for people to create fake profiles for defaming accounts they dislike. This type of users creates accounts with the username of the people they hate and post irrelevant stories and snap shots on their accounts to redirect everybody so that they assume that particular person is awful and make their reputation low.

Most attackers are in it to make money [3]. They make money by distributing unwanted ads (spam) or capturing accounts they can reuse or resale (phishing). Spammers gather resources to know fake and real users, email ids, Ip locations and computing knowledge power. Every one of these advantages can have a huge expense related with them, and an assault, similar to any business adventure, needs benefit to continue onward. Attackers more often use Facebook logins, applications, Events, Group users to gather login credentials, spam users, and ultimately gain profits [4]. They need email records, treats, and a wide scope of IP delivers to go around notoriety-based protections. Moreover, they use telephone numbers, taken charge cards, and CAPTCHA arrangements trying to go around validation checks.

Facebook security privileges its system to gather users to prevent spams and fishing accounts [5]. Facebook Immune System does continuous minds all gather and each its activity made by it. Social bot is a known that stops and controls social online accounts. Bots socially is an auto generated software. Precised way a social account duplicates relies upon at the social media, also in contrast to general bot, a social bot interacting more in different customers that the social bot is an actual man or woman.

2. Literature Survey

Wanda, et al. [6] proposed finding abnormal nodes in online social networks using dynamic deep learning. The authors propose a model to classify malicious vertices using nodes' link information by training extensive features with dynamic deep learning architecture. Initially, to construct dynamic deep learning, they present a generic function called WalkPool pooling to optimize their network performance. By demonstrating the proposed model, they gain higher accuracy than standard learning algorithms in the abnormal nodes' classification.

Senthil raja, et al. [7] proposed detection of malicious profiles and protecting users in online social networks. Initially, by the assessment of 3PS (Publicly Privacy Protection System), this work employs the malicious account detection method in OSN depending upon the mischievous person's uncountable shared posts in a day and latest activity and behaviors. Examining the network similarity and comparison of attributes threshold values referred to the original user's profile can be used to identify the malicious accounts. For this E-SVM-NN classifier is used based on the feature reduction techniques. This work involves in creating OSN accounts for experiments and investigates the latest updates, posts, comments, photos, and performing online search etc. which are used to evaluate the effectiveness and significance of the proposed work in contrast to the previous works.

Bhattacharya, et al. [8] proposed Application of Machine Learning (ML) Techniques in Detecting Fake Profiles on social media. Besides the wide range of advantages that social media offers, it also comes up with many disadvantages of being an online platform. Issues like fake profiles and impersonation have increased on social media platforms, such as computer-generated bots, human-generated, or cyborgs. Such accounts are made with malicious intentions. Moreover, there is no feasible solution to such a problem. They developed a model which can detect fake profiles on social media by using ML techniques for better prediction and identification. Instagram data is considered for the availability of the dataset, and the analysis will be done by implementing the ML algorithms, which gives the highest accuracy on the dataset. By accurately detecting such profiles, social media platforms can be made safe for users.

Singh, et al. [9] proposed Predicting image credibility in fake news over social media using multi-modal approach. Fake images are often associated with textual data. Hence, a multi-modal framework is employed utilizing visual and textual feature learning. However, few multi-modal frameworks are already proposed; they are further dependent on additional tasks to learn the correlation between modalities. An efficient multi-modal approach is proposed, which detects fake images of microblogging platforms. No further additional subcomponents are required. The proposed framework utilizes explicit convolution neural network model EfficientNetB0 for images and sentence transformer for text analysis. The feature embedding from visual, and text is passed through dense layers and later fused to predict fake images.

Xu, et al. [10] proposed Deep entity classification: Abusive account detection for online social networks. However, a practical, effective ML-based defense requires carefully engineering features that are robust to adversarial manipulation, obtaining enough ground truth labeled data for model training, and designing a system that can scale to all active accounts on an OSN (potentially in the billions). To address these challenges, they present Deep Entity Classification (DEC), an ML framework that detects abusive accounts in OSNs that have evaded other, traditional abuse detection systems.

Karami, et al. [11] proposed Profiling Fake News Spreaders on social media through Psychological and Motivational Factors. A majority of methods developed to combat disinformation either focus on fake news content or malicious actors who generate it. However, the virality of fake news is largely dependent upon the users who propagate it. A deeper understanding of these users can contribute to the development of a framework for identifying users who are likely to spread fake news. In this work, they study the characteristics and motivational factors of fake news spreaders on social media with input from psychological theories and behavioral studies. They then perform a series of experiments to determine if fake news spreaders can be found to exhibit different characteristics than other users.

Khanday, et al. [12] proposed Artificial Neural Network-Based Propaganda Identification on social media in COVID-19 Era. Initially, the data are extracted using multiple ambiguous hashtags and are manually annotated into binary class. Hybrid feature engineering is being performed by combining “Term Frequency (TF)/Inverse Document Frequency (IDF),” “Bag of Words,” and Tweet Length. The proposed algorithm is compared with logistic regression, support vector machine, and multinomial Naive Bayes.

2.1 MOTIVATION

As the number of people using OSN increases, so does the fake social media accounts creation. The main motivational factor in identifying those fake accounts is the cyber-crime rate, as these accounts were created primarily to commit cyber robbery or to commit cybercrime anonymously or unidentified is a significant increase from last few years. Fake account owners also try to take advantage of people's kindness by composing fake messages and spreading false news through these fake accounts to steal money from sinless people.

In addition, people want to create multiple accounts that don't belong to anyone, created just to raise votes in an online voting system, and receive referral incentives, as in online games. The detection of fake accounts in OSN attracts many researchers, so several algorithms for detection of fake accounts have been developed using ML techniques and various functions to connect to the account. Spammers can also find ways to support such techniques. These security technologies provide sophisticated detection mechanisms that require the continuous development of new approaches to spam detection. The main hazards in detection of fake accounts are to achieve accuracy and response time characteristics.

3. Proposed System

Online social networks such as Facebook or Twitter contains user’s details, and some malicious users will hack social network database to steal or breach users’ information. Therefore, to protect the users’ data, this project uses artificial neural networks (ANNs) to identify whether given social network account details are from genuine or fake users. ANN algorithm will be trained with all previous users fake and genuine account dataset and then whenever we gave new test data then ANN pretrained model will be applied on new test data to identify whether the given new account details are from genuine or fake users. Usually, all fake user's main intention is to send friend request to normal users to hack their machine or to steal their data and never they will have many numbers of posts or have many following friends and their account age also will have a smaller number of years. By analysing this features Facebook will mark whether user profile is fake or genuine.

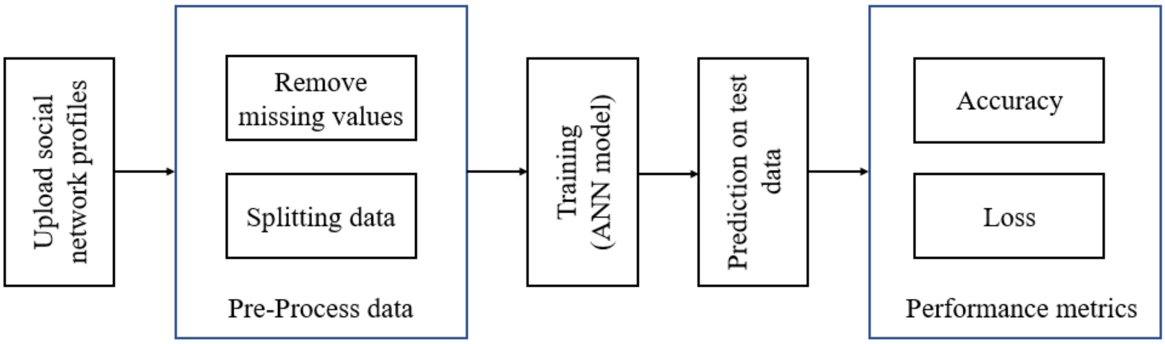


Fig. 1: Proposed architecture of fake profile identification using ANN.

3.1 What is a Neural Network?

The structure of the human brain inspires a neural network. It is a web of interconnected entities known as nodes wherein each node is responsible for a simple computation. In this way, a Neural Network functions similarly to the neurons in the human brain. They connect various nodes, and each node is tasked with a direct computation. In other words, neural networks represent a collection of algorithms developed to identify patterns. They classify or group raw input to comprehend sensory data through machine perception. The real-world data, like texts, images, sounds, etc., should be transformed into vectors to allow neural networks to identify the patterns. Neural networks can't learn the fundamental rules if your dataset is small. Its internal structure is difficult to comprehend.

The following benefits help you to compare neural networks and deep learning.

- Stores information on the entire network: In conventional programming, data is stored on the network instead of a database. The neural networks make sure the entire network's operation is not stopped when a few pieces of data disappear from a location. They provide good fault tolerance. They make sure the corruption of one or multiple artificial network cells doesn't impact the output production. Hence, networks can better tolerate errors.
- Distributed memory: Two aspects are important to allow an artificial neural network to learn. They outline the examples and train the network as per the anticipated output by offering related examples. These examples are directly correlated with the network's development. neural networks and deep learning
- Can work with incomplete knowledge: The output produced by the data may be incomplete. The neural networks can work on this data to identify the missing aspect and work accordingly.
- Avoids network corruption: A network can slow down or degrade over time. The neural networks protect the data from this corruption.
- Trains a machine: An artificial neural network can comment on comparable situations. Consequently, they can learn from these experiences and make decisions.

- Supports parallel processing: Their ability to parallel process helps them to accomplish multiple tasks simultaneously. It is one of the prominent aspects that differentiates deep learning neural networks.

ADVANTAGES

- ANN arranges algorithms in a fashion that it can make accurate decisions by itself.
- They do not require human intervention as the nested layers within pass the data through hierarchies of various concepts, which eventually makes them capable of learning through their own errors.

3.2 Proposed ANN algorithms details

To demonstrate how to build ANN neural network-based classifier, we shall build a 6-layer neural network that will identify and separate one from other. This network that we shall build is a very small network that we can run on a CPU as well. Traditional neural networks that are very good at doing fake profile identification have many more parameters and take a lot of time if trained on normal CPU. However, our objective is to show how to build a real-world neural network using TENSORFLOW.

Neural Networks are essentially mathematical models to solve an optimization problem. They are made of neurons, the basic computation unit of neural networks. A neuron takes an input (say x), do some computation on it (say: multiply it with a variable w and adds another variable b) to produce a value (say; z= wx + b). This value is passed to a non-linear function called activation function (f) to produce the final output (activation) of a neuron. There are many kinds of activation functions. One of the popular activation functions is Sigmoid. The neuron which uses sigmoid function as an activation function will be called sigmoid neuron. Depending on the activation functions, neurons are named and there are many kinds of them like RELU, TanH. If you stack neurons in a single line, it’s called a layer; which is the next building block of neural networks. See below image with layers.

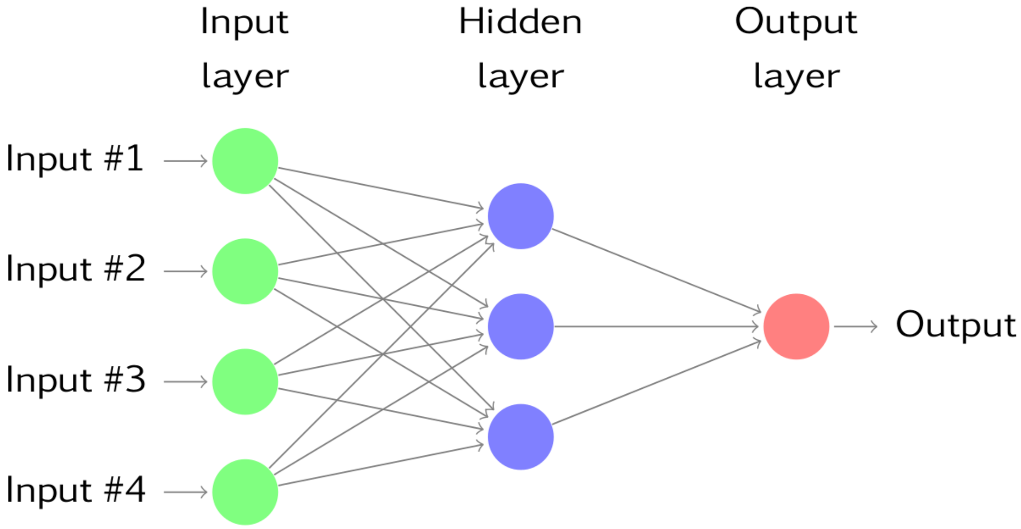


Fig. 2: Proposed ANN algorithm.

To predict class label, multiple layers operate on each other to get best match layer and this process continues till no more improvement left.

3.3 MODULE IMLEMENTATION

Upload Social Network Profiles Dataset: Using this module we will upload dataset to application

Pre-process Dataset: Using this module we will apply processing technique such as removing missing values and then split dataset into train and test where application use 80% dataset to train ANN and 20% dataset to test ANN prediction accuracy

Run ANN Algorithm: Using this module we will train ANN algorithm with train and test data and then train model will be generated and we can use this train model to predict fake accounts from new dataset.

ANN Accuracy & Loss Graph: To train ANN model we are taking 200 epoch/iterations and then in graph we will plot accuracy/loss performance of ANN at each epoch/iteration.

Predict Fake/Genuine Profile using ANN: using this module we will upload new test data and then apply ANN train model to predict whether test data is genuine or fake.

4. Results and Discussion

4.1 Dataset Description

To train ANN algorithm we are using below details from social networks.

4.1.1 Account_Age, Gender, User_Age, Link_Desc, Status_Count, Friend_Count, Location, Location_IP, Status

All fake user’s main intention is to send friend request to normal users to hack their machine or to steal their data and never they will have many number of posts or have many following friends and their account age also will have less number of years. By analysing this features Facebook will mark whether user profile is fake or genuine. This Facebook profile data we downloaded from Facebook website and using this data to train ANN model. Below are some values from profile dataset.

4.1.2 Account_Age, Gender, User_Age, Link_Desc, Status_Count, Friend_Count, Location, Location_IP, Status

10, 1, 22, 0, 1073, 237, 0, 0, 0

10, 0, 33, 0, 127, 152, 0, 0, 0

10, 1, 46, 0, 1601, 405, 0, 0, 0

10, 0, 25, 0, 704, 380, 0, 0, 0

7, 1, 34, 1, 64, 721, 1, 1, 1

7, 1, 30, 1, 69, 587, 1, 1, 1

7, 1, 36, 1, 61, 782, 1, 1, 1

7, 1, 52, 1, 96, 827, 1, 1, 1

In above dataset all bold names are the dataset column names and all integer values are the dataset values. As ANN will not take string value so we convert gender values to 0 or 1 if male value is 1 and if female value is 0. In above dataset last column give us information of fake or genuine account if last column contains value 0 then account is genuine otherwise fake. All fake account will have a smaller number of posts as their main intention is to send friend requests do not post, so by analysing this features Facebook mark that record with value 1 which means it’s a fake account.

Below are some values from test data

4.1.3 Account_Age, Gender, User_Age, Link_Desc, Status_Count, Friend_Count, Location, Location_IP

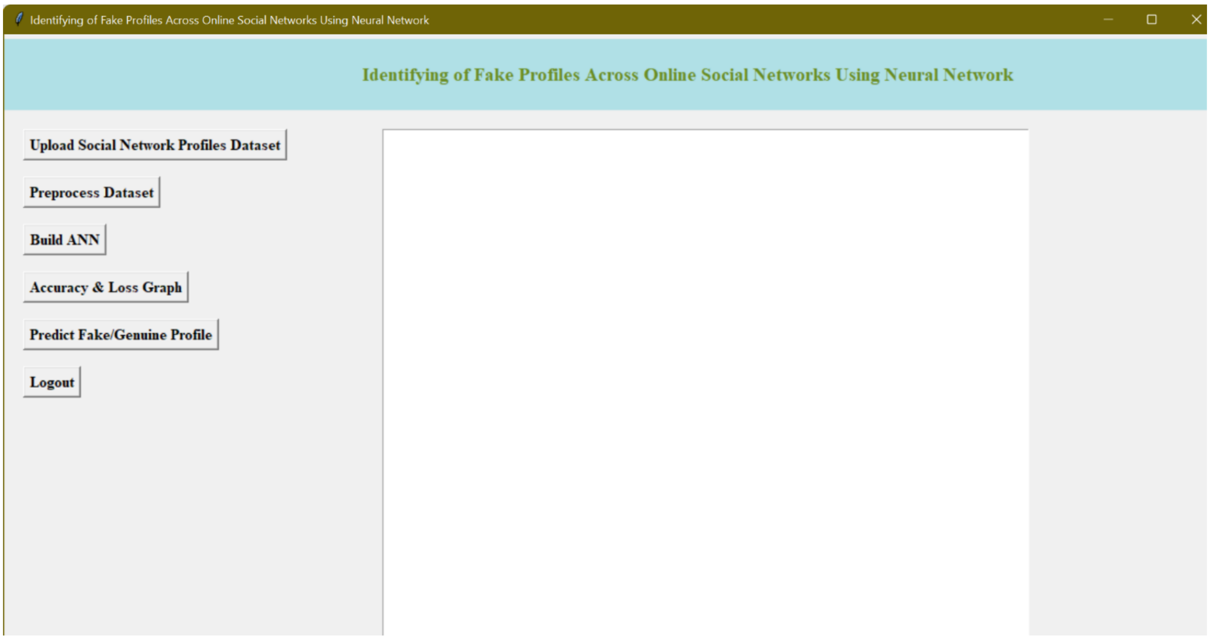
10, 1, 44, 0, 280, 1273, 0, 0

10, 0, 54, 0, 5237, 241, 0, 0

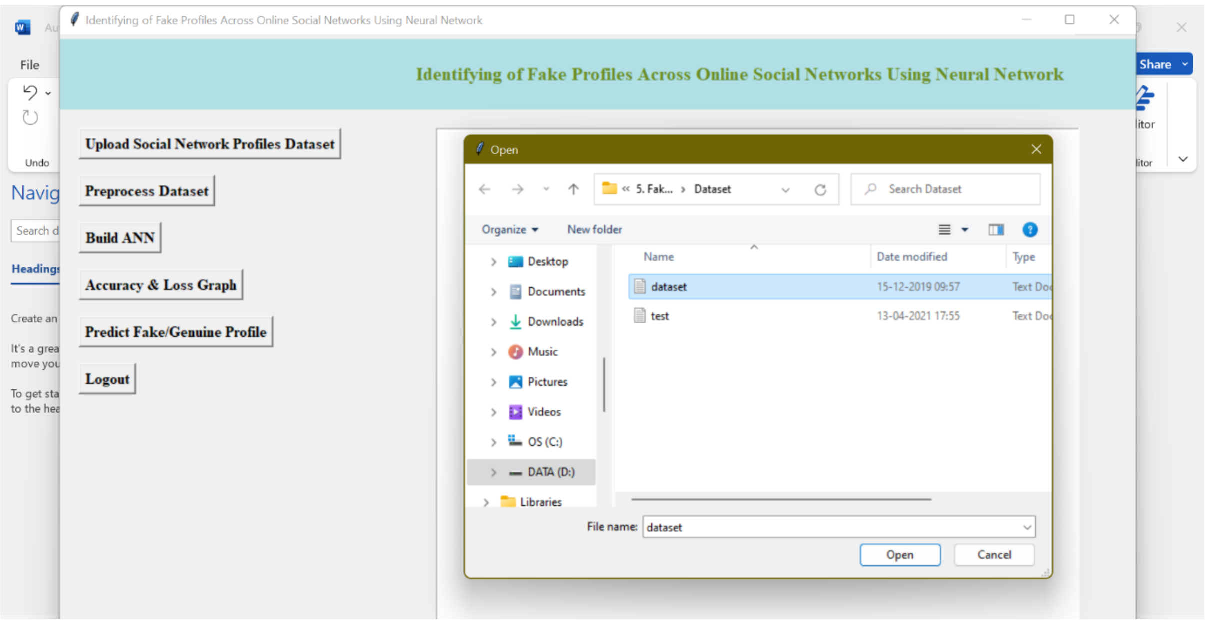
7, 0, 42, 1, 57, 631, 1, 1

7, 1, 56, 1, 66, 623, 1, 1

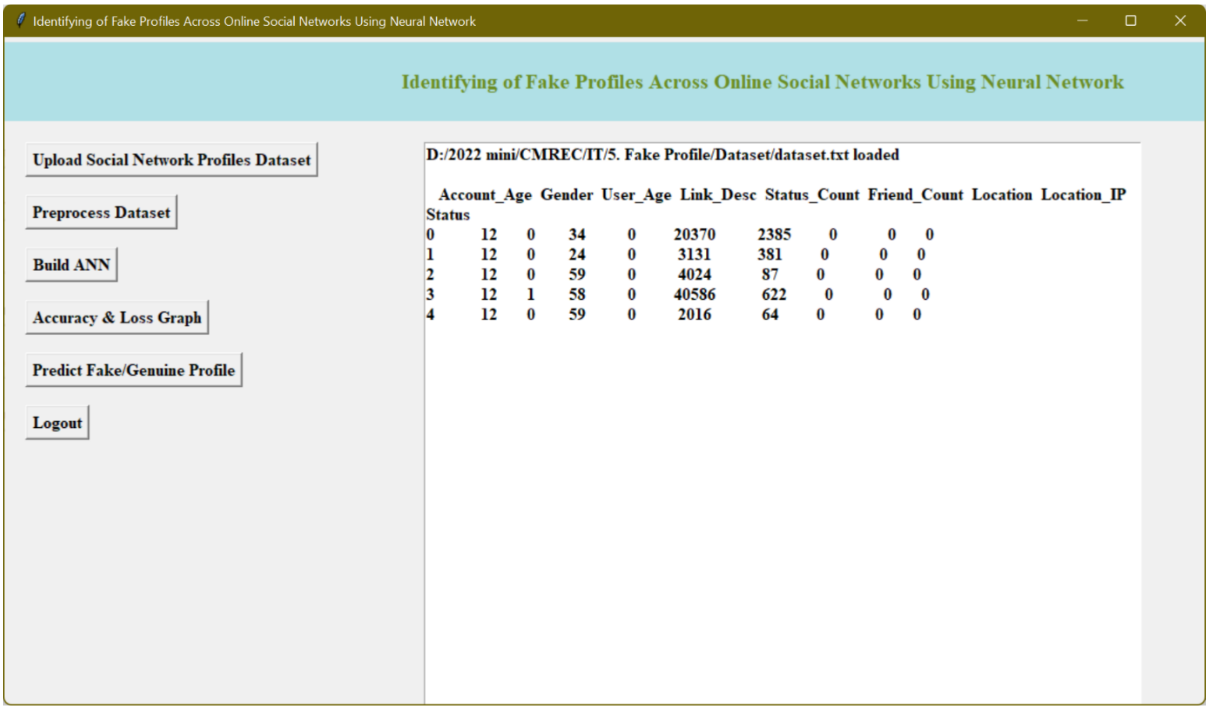
In above test data STATUS column and its value is not there and ANN will predict status and give us result whether above test data is fake or genuine.


Identifying of Fake Profiles Across Online Social Networks Using Neural Network
Identifying of Fake Profiles Across Online Social Networks Using Neural Network
Upload Social Network Profiles Dataset
Preprocess Dataset
Build ANN
Accuracy & Loss Graph
Predict Fake/Genuine Profile
Logout

In above screen click on 'Upload Social Network Profiles Dataset' button and upload dataset



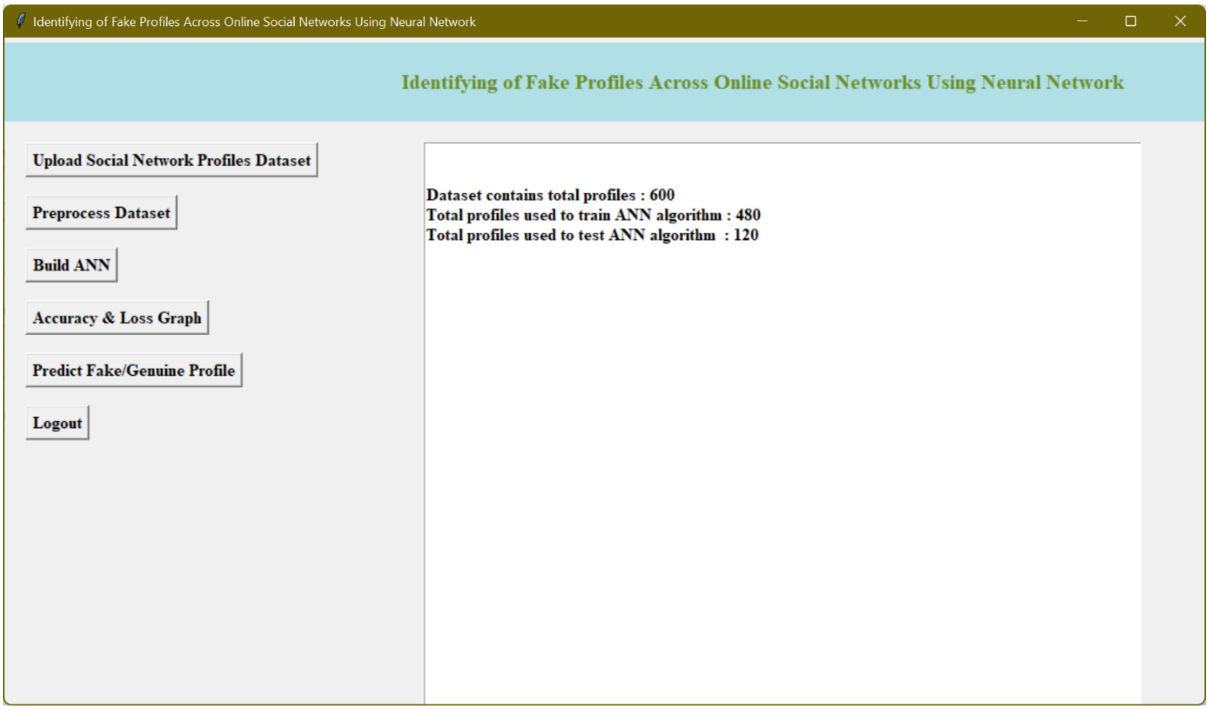
In above screen selecting and uploading 'dataset.txt' file and then click on 'Open' button to load dataset and to get below screen


Identifying of Fake Profiles Across Online Social Networks Using Neural Network
Identifying of Fake Profiles Across Online Social Networks Using Neural Network
Upload Social Network Profiles Dataset
Preprocess Dataset
Build ANN
Accuracy & Loss Graph
Predict Fake/Genuine Profile
Logout

```
D:/2022 mini/CMREC/IT/5. Fake Profile/Dataset/dataset.txt loaded

  Account_Age  Gender  User_Age  Link_Desc  Status_Count  Friend_Count  Location  Location_IP
Status
0        12      0       34         0       20370        2385       0         0     0
1        12      0       24         0        3131         381      0         0     0
2        12      0       59         0        4024          87     0         0     0
3        12      1       58         0       40586         622      0         0     0
4        12      0       59         0        2016          64     0         0     0
```

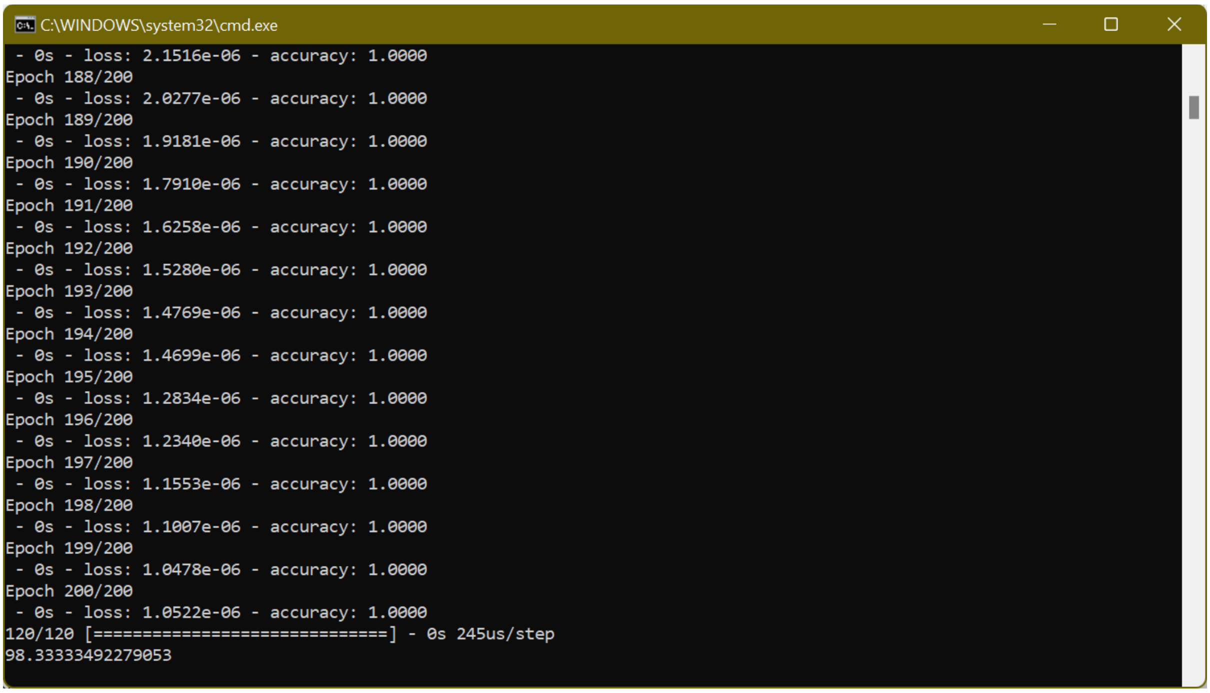
In above screen dataset loaded and displaying few records from dataset and now click on 'Preprocess Dataset' button to remove missing values and to split dataset into train and test part


Identifying of Fake Profiles Across Online Social Networks Using Neural Network
Identifying of Fake Profiles Across Online Social Networks Using Neural Network
Upload Social Network Profiles Dataset
Preprocess Dataset
Build ANN
Accuracy & Loss Graph
Predict Fake/Genuine Profile
Logout

```
Dataset contains total profiles : 600
Total profiles used to train ANN algorithm : 480
Total profiles used to test ANN algorithm  : 120
```

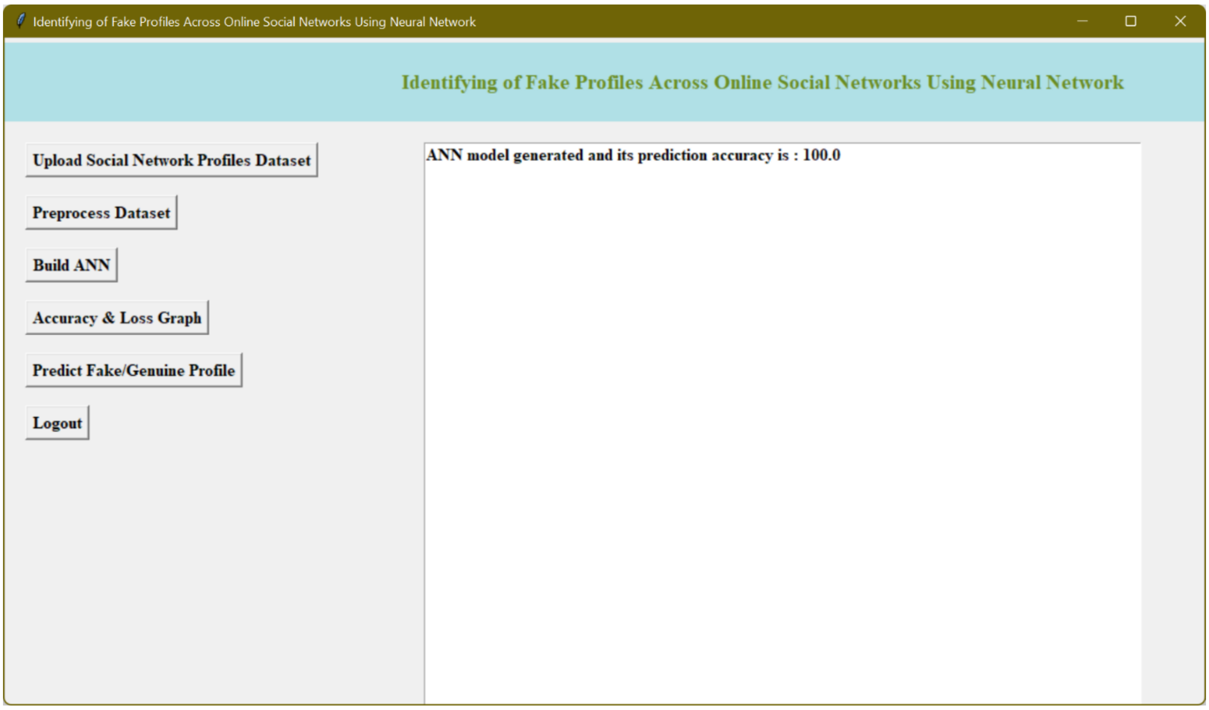
In above screen we can see dataset contains total 600 records and application using 480 records for training and 120 records to test ANN and now dataset is ready and now click on 'Run ANN Algorithm' button to ANN algorithm

In below screen we can see ANN start iterating model generation and at each increasing epoch we can see accuracy is getting increase and loss getting decrease.



```
 - 0s - loss: 2.1516e-06 - accuracy: 1.0000
Epoch 188/200
 - 0s - loss: 2.0277e-06 - accuracy: 1.0000
Epoch 189/200
 - 0s - loss: 1.9181e-06 - accuracy: 1.0000
Epoch 190/200
 - 0s - loss: 1.7910e-06 - accuracy: 1.0000
Epoch 191/200
 - 0s - loss: 1.6258e-06 - accuracy: 1.0000
Epoch 192/200
 - 0s - loss: 1.5280e-06 - accuracy: 1.0000
Epoch 193/200
 - 0s - loss: 1.4769e-06 - accuracy: 1.0000
Epoch 194/200
 - 0s - loss: 1.4699e-06 - accuracy: 1.0000
Epoch 195/200
 - 0s - loss: 1.2834e-06 - accuracy: 1.0000
Epoch 196/200
 - 0s - loss: 1.2340e-06 - accuracy: 1.0000
Epoch 197/200
 - 0s - loss: 1.1553e-06 - accuracy: 1.0000
Epoch 198/200
 - 0s - loss: 1.1007e-06 - accuracy: 1.0000
Epoch 199/200
 - 0s - loss: 1.0478e-06 - accuracy: 1.0000
Epoch 200/200
 - 0s - loss: 1.0522e-06 - accuracy: 1.0000
120/120 [==============================] - 0s 245us/step
98.33333492279053
```

In above screen we can see after 200 epoch ANN got 98.33% accuracy and in below screen we can see final ANN accuracy


Identifying of Fake Profiles Across Online Social Networks Using Neural Network
Identifying of Fake Profiles Across Online Social Networks Using Neural Network
Upload Social Network Profiles Dataset
Preprocess Dataset
Build ANN
Accuracy & Loss Graph
Predict Fake/Genuine Profile
Logout

ANN model generated and its prediction accuracy is : 100.0

In above screen ANN model generated and now click on ‘ANN Accuracy & Loss Graph’ button to get below graph

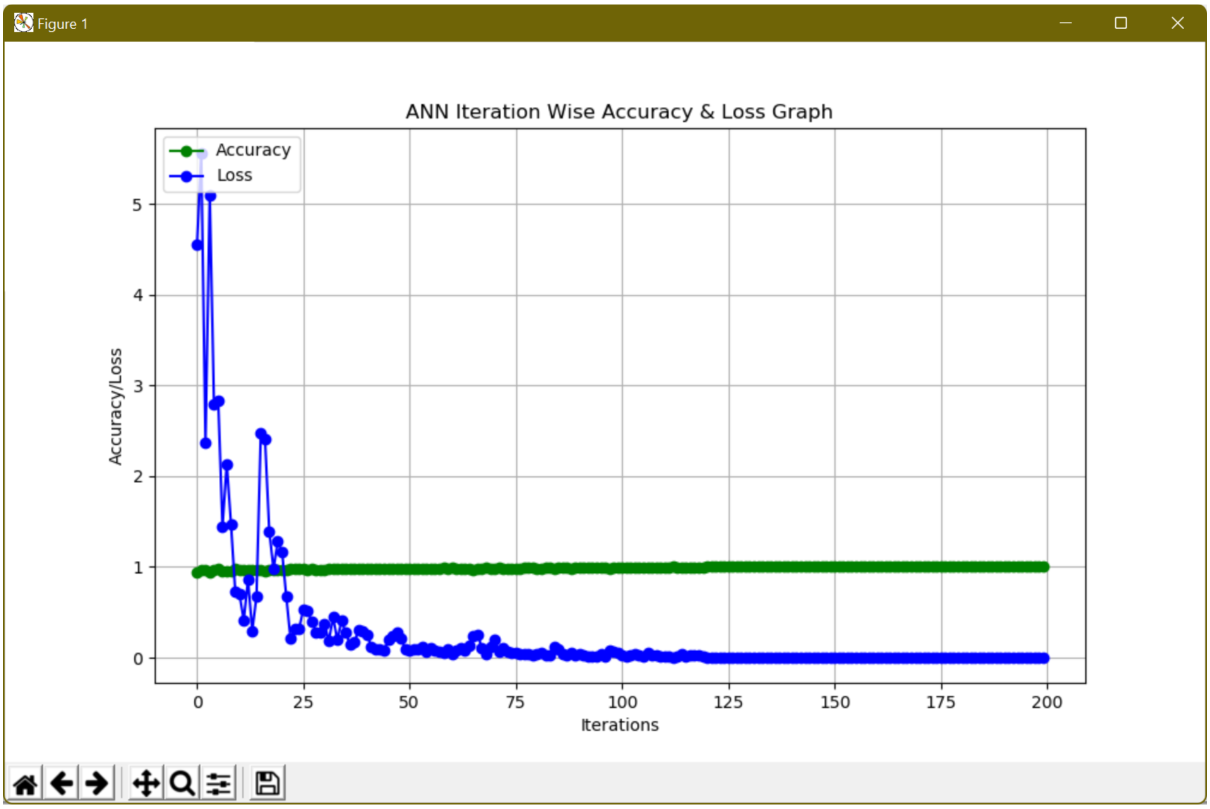
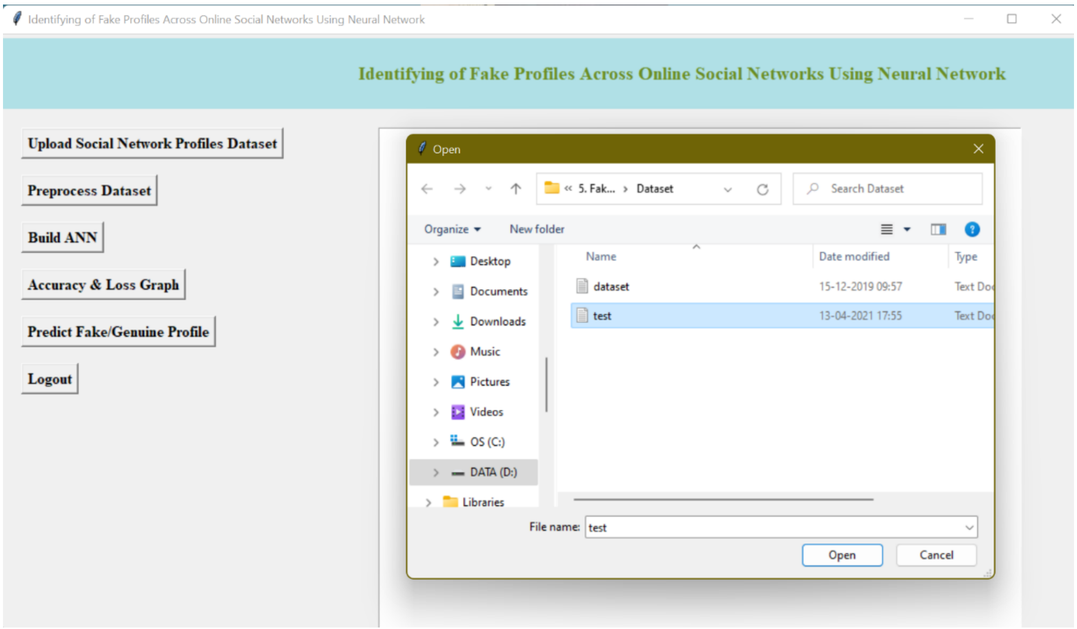


Figure 3. ANN iteration Wise Accuracy & Loos Graph

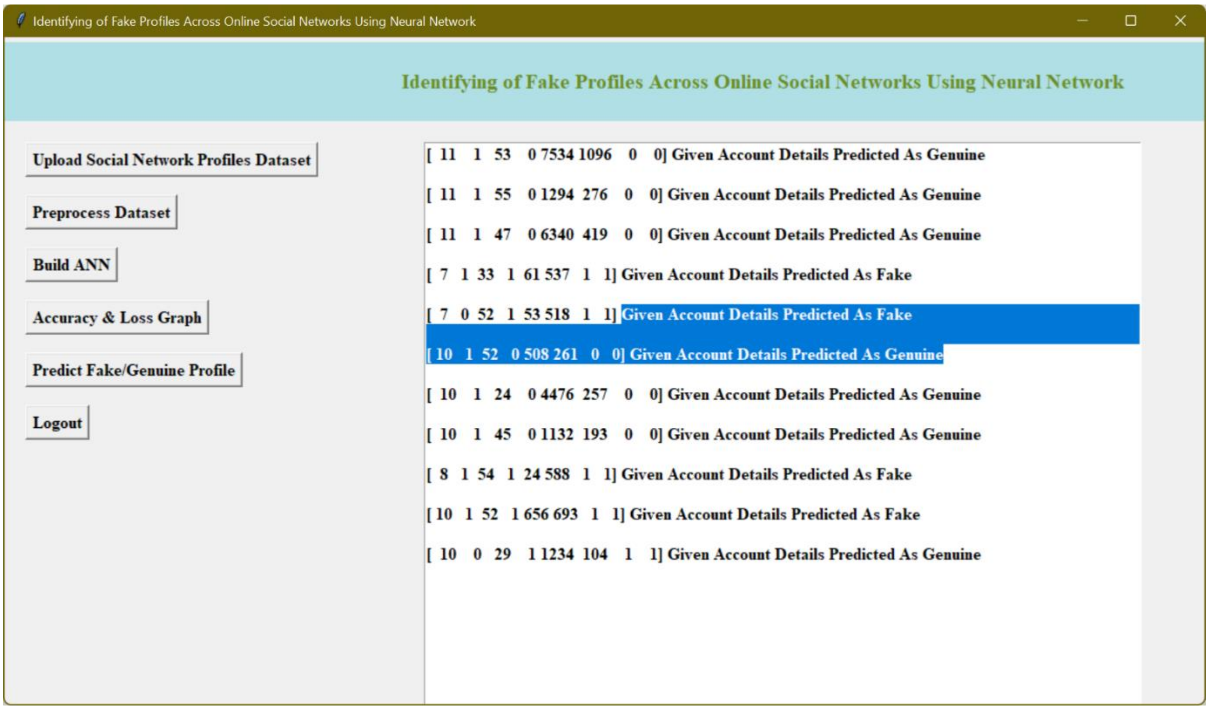
In above graph x-axis represents epoch and y-axis represents accuracy/loss value and in above graph green line represents accuracy and blue line represents loss value and we can see accuracy was increase from 0.90 to 1 and loss value decrease from 7 to 0.1. Now model is ready and now click on 'Predict Fake/Genuine Profile using ANN' button to upload test data and then ANN will predict below result


Identifying of Fake Profiles Across Online Social Networks Using Neural Network
Identifying of Fake Profiles Across Online Social Networks Using Neural Network
Upload Social Network Profiles Dataset
Preprocess Dataset
Build ANN
Accuracy & Loss Graph
Predict Fake/Genuine Profile
Logout
Open
« 5. Fak... › Dataset
Search Dataset
Organize
New folder
Desktop
Documents
Downloads
Music
Pictures
Videos
OS (C:)
DATA (D:)
Libraries
Open
Cancel

Name	Date modified	Type
dataset	15-12-2019 09:57	Text Doc
test	13-04-2021 17:55	Text Doc

File name: test

In above screen we are selecting and uploading 'test.txt' file and then click on 'Open' button to load test data and to get below prediction result


Identifying of Fake Profiles Across Online Social Networks Using Neural Network
Identifying of Fake Profiles Across Online Social Networks Using Neural Network
Upload Social Network Profiles Dataset
Preprocess Dataset
Build ANN
Accuracy & Loss Graph
Predict Fake/Genuine Profile
Logout

```
[ 11   1  53   0 7534 1096   0   0] Given Account Details Predicted As Genuine
[ 11   1  55   0 1294  276   0   0] Given Account Details Predicted As Genuine
[ 11   1  47   0 6340  419   0   0] Given Account Details Predicted As Genuine
[ 7  1  33  1  61 537  1  1] Given Account Details Predicted As Fake
[ 7  0  52  1  53 518  1  1] Given Account Details Predicted As Fake
[ 10  1  52  0  508 261  0  0] Given Account Details Predicted As Genuine
[ 10   1  24   0 4476  257   0   0] Given Account Details Predicted As Genuine
[ 10   1  45   0 1132  193   0   0] Given Account Details Predicted As Genuine
[ 8  1  54  1  24 588  1  1] Given Account Details Predicted As Fake
[ 10  1  52  1  656 693  1  1] Given Account Details Predicted As Fake
[ 10   0  29   1 1234  104   1   1] Given Account Details Predicted As Genuine
```

In above screen in square bracket, we can see uploaded test data and after square bracket we can see ANN prediction result as genuine or fake.

5. Conclusion

Online social networks such as Facebook or Twitter contains user's details, and some malicious users will hack social network database to steal or breach users' information. Therefore, this work uses ANN to identify whether given social network account details are from genuine or fake users. ANN algorithm will be trained with all previous users fake and genuine account dataset and then whenever a new test data is given for prediction then the trained model will be applied on new test data to identify whether given new account details are from genuine or fake users.

REFERENCES

[1] Singh, Vernika, Raju Shanmugam, and Saatvik Awasthi. "Preventing Fake Accounts on Social Media Using Face Recognition Based on Convolutional Neural Network." Sustainable Communication Networks and Application. Springer, Singapore, 2021. 227-241.

[2] Sahoo, Somya Ranjan, and Brij B. Gupta. "Multiple features-based approach for automatic fake news detection on social networks using deep learning." Applied Soft Computing 100 (2021): 106983.

[3] Awan, Mazhar Javed, et al. "Fake profile recognition using big data analytics in social media platforms." International Journal of Computer Applications in Technology 68.3 (2022): 215-222.

[4] Prabhu Kavin, B., et al. "Machine learning-based secure data acquisition for fake accounts detection in future mobile communication networks." Wireless Communications and Mobile Computing 2022 (2022).

[5] Mujeeb, Shaik, and Sangeeta Gupta. "Fake Account Detection in Social Media Using Big Data Analytics." Proceedings of Second International Conference on Advances in Computer Engineering and Communication Systems. Springer, Singapore, 2022.

[6] Wanda, Putra, and Huang J. Jie. "DeepFriend: finding abnormal nodes in online social networks using dynamic deep learning." Social Network Analysis and Mining 11.1 (2021): 1-12.

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

[8] Bhattacharya, Ananya, et al. "Application of Machine Learning Techniques in Detecting Fake Profiles on Social Media." 2021 9th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO). IEEE, 2021.

[9] Singh, Bhuvanesh, and Dilip Kumar Sharma. "Predicting image credibility in fake news over social media using multi-modal approach." Neural Computing and Applications (2021): 1-15.

[10] Xu, Teng, et al. "Deep entity classification: Abusive account detection for online social networks." 30th {USENIX} Security Symposium ({USENIX} Security 21). 2021.

[11] 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.

[12] Khanday, Akib Mohi Ud Din, et al. "NNPCov19: Artificial Neural Network-Based Propaganda Identification on Social Media in COVID-19 Era." Mobile Information Systems 2022 (2022).