Detecting Fake Social Media Account Using Deep Neural Networking

Mansi Kesharwani[1], Surbhi Kumari[2], Vandana Niranjan[3]

[1,2]Student, Dept. of Electronics and Communication Engineering, Indira Gandhi Delhi Technical University for Women, Delhi, India
[3]Professor, Dept. of Electronics and Communication Engineering, Indira Gandhi Delhi Technical University for Women, Delhi, India country

---***---

Abstract - *Social media is presently a significant piece of our daily life. Currently more than the half of the world is an active user of the social media platforms. The ever - increasing popularity of these platforms has also given rise to a major issue which is the presence of fake accounts on them. These fake accounts serve the purpose of impersonating or cat-fishing other people. They have become an easy way to sell fake products and services to the customers. Also, the personal data of billions of people are at stake. These threats have made it essential to detect and deactivate the dummy accounts before any harm gets done. By the virtue of Machine Learning it has become easy to automatically detect millions of such accounts in a matter of seconds. In this project, we explore a deep learning model that can be used to classify a given account as real or fake, especially in Instagram. In the proposed work accuracy of the model is 93.63 percent.*

Key Words: Social media, Instagram, Deep Neural Network, Artificial Neural Network, Fake account

1.INTRODUCTION

Throughout the human history people have worked on developing better ways of communication. One such attempt led to the birth of social networking sites in the 1990s. Though it took sometime for people to catch on with these sites but by the early 2000s the social media platform began to flourish [1]. At least 3 social networking site were launched every year. The world witnessed one of the biggest revolution of all time.

Currently everyone feels the need to have an online presence. The social media has become the number 1 activity on the internet [2]. In 2010 the number of active users of social media was 970 Million, within a decade this number rose upto 3.6 Billion in 2020. In the coming 5 years it is expected that the social media will continue to bloom, eventually attracting more than 4.41 Billion users by the year 2025.

These platforms have transformed the web into a social web where it has become easy for people to find old friends via Facebook, get latest news updates on Twitter [3], job hunt on LinkedIn and watching trending videos with a single click on Youtube. These platforms have become our go to space for entertainment and keeping abreast. Not only this social media has also given us a space to voice our opinion without any fear. As a result, we have witnessed hundreds of revolutions on the internet like the popular me-too movement. These platforms have also evolved into a new marketing platform which is free of cost unlike the traditional forms of publicity like TV, billboards, radio etc. Social media has become so entangled with our everyday activities that it has become impossible to imagine our lives without these platforms.

The growing popularity of social media platforms has not only benefitted the people but also caught the attention of scammers. On one hand social media is bringing people together and on the other hand it has created a guarded space for fraudsters to carry out a number of illegal activities. The absence of any authentication process has made it easy for anyone to make a fake account. This serves as an advantage for the scammers encouraging them to use fake account for illegal activities as there is a good chance that the account holder will not get caught. Owing to this the popularity of fake accounts has increased. These phoney accounts can either operated by humans or bots.

The use of these phantom accounts to impersonate someone in hope of defaming them has become a common issue. At times these accounts serve the bigger purpose of acting as a trusted acquaintance to get personal information from a person. This obtained information can be used to carry out phishing attacks.

People often use these dummy accounts to spread fake news which in the worst case can cause riot like conditions. Some people make use of fake accounts to spread hate which can be directed at certain race, religion, country or often at a particular person. This has increased the cases of cyber bullying leading to rise in the cases of depression and anxiety in teenagers.

The social media platforms have also seen an increase in the number of accounts which provide services or products in exchange of money. But most of these accounts are fake, as a result thousands of people are sold fake products and are promised fake services by these accounts. Sometimes these fake accounts are used by companies to build hype for their bad products and services [4]. Not only scammers but also a lot of influencers also use fake bot followers to appear popular, which helps them in gaining more offers from companies asking to publicise their products. At times the fake accounts can also be used to befriend a person in order to stalk them.

Another big issue associated with the fake accounts is the amount of data overload that they are resulting in. With such a large number of users the social media platforms create a lot of data. It is estimated that by the year 2025 even Google data centres will suffer from data overload. According to a survey 500 Million women have Facebook account but the population of women in the world is only 300 Million [5]. The survey also reveals similar case for other social media platforms as well.

The growing threats of these fake accounts has made it necessary to take them down. With the number of fake accounts being in millions it has become impossible to manually detect them. Luckily the advancement in digital technology can benefit a lot in this situation. Methods like Machine Learning can help in making the stratification process a lot easier and accurate.

This project involves use of deep learning model to classify social media accounts as genuine or fake. An Artificial Neural Network (ANN) model is used to support the stratification process. The upshot of the model is analysed using confusion matrix and graphs.

2. LITERATURE SURVEY

Researchers around the globe have worked on a lot of methodologies to detect phantom profiles. These method involve extracting some features from the target account and feeding them to a trained Machine Learning classifier. This classifier looks for patterns in the given data and eventually based on the discovered pattern labels a given account as legitimate or fake.

For instance Samala Durga Prasad Reddy (2019) [6] used a random forest classifier to detect fake accounts with 95% accuracy. Profile features like id, name, status count, followers count, friends count, location, date of creation of the id, numbers of shares done by the account, gender and language used by the account holder were used as features for the classification process.

Rohit Raturi (2018) [5] proposed 2 architectures for solving this issue. The first one uses NLP and it marks 2 or more accounts as suspicious if they use the same IP or MAC address. In the second architecture Support Vector Machine(SVM) is used for finding out accounts which make frequent use of harmful words. These suspicious accounts then have to verify themselves.

Yasyn Elyusufi, Zakaria Elyusufi, and M'hamed Ait Kbir [7] built 3 supervised machine learning algorithms models using naïve bayes with 78.33% accuracy, decision tree with 99.28% accuracy and random forest with 99.64% accuracy to distinguish real and fake Facebook accounts.

S. P. Maniraj, Harie Krishnan G, Surya T, Pranav R (2019)[8] have defined a profile as phoney on the off chance that it doesn't coordinate with the characterised commitment rate or has actorly activity or is involved in spamming. They have utilised gradient boosting alongside decision tree for the purpose.

The researches done so far make use of traditional machine learning algorithms like random forest, naive bayes, SVM, and decision tree. These methods are incapable of doing feature selection on their own. Thus the researcher has to study relation between the features and target variable in order to decide which features are to be considered and which can be rejected. Another drawback being their inability to adapt with the changing patterns in the input dataset which can make them insufficient at times. Hence these methods require constant monitoring. The changing patterns in the input can cause them to give incorrect results thus reducing the accuracy. Also one major issue with them is that they do not perform well if the dataset is too large or is unstructured. This makes traditional methods highly unsuitable for real life scenarios as in such cases the data is mostly unstructured and often too large.

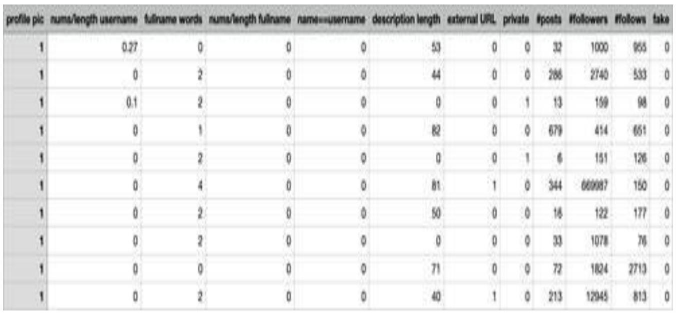
Owing to the drawbacks of the traditional methods it has become necessary to explore advance algorithms like deep neural network. Deep neural network is a highly efficient method which can be used to replace the traditional machine learning methods. It works very well on unstructured data and perfectly deals with large datasets. The major drawback of feature engineering which existed in case of traditional machine learning algorithms is eliminated in case of deep neural network. It does not require the researcher to specify the features to be used instead it itself decides which features are important and which do not affect the result. This makes deep learning effective and accurate.

3. PROPOSED SOLUTION

Each social media profile has a lot of data associated with it like the user name, account holder's name, number of friends of the user, date of birth of the user, phone number of the user, etc. We can make use of the associated data to comment on the genuineness of the account. In this research we have used deep neural network to recognise fake Instagram accounts. A six layered ANN model is used for the purpose. The designed model uses features like username, number of followers, profile description length, number of accounts followed by the user, number of posts etc to declare a given account as genuine or fake. Nowadays for security reasons the social media companies have started providing the facility of making account private to the users but if this method is deployed then they can even access the information of private accounts for examination without any violation of privacy [6].

3.1 Dataset Used

The dataset used for this case study is an open dataset which is available on Kaggle. It was collected using a crawler by Bardiya Bakhshandeh [9]. The dataset has details of 696 Instagram accounts. These 696 data entries are divided into 2 folders train set and test set. The train set has data of 576 accounts and the test set has data of 120 accounts. Both train and test set are balanced meaning they have a ratio of 1:1 between real and fake accounts, i.e., the train set has 288 real and 288 fake account details and test set has 60 real and 60 fake account entires. The target variable which expresses whether the given account is real or fake can take up 2 values 0 and 1. The value being 1 if the account is fake and 0 if the account is genuine. Other than the target variable the dataset has 10 features out of which 4 take binary value and the rest take integer value. Before the dataset can be used for the estimation process it is essential to process it. The dataset was checked in order to eliminate any missing values. Fig-1 shows the dataset used for the training purpose.



profile pic	nums/length username	fullname words	nums/length fullname	name==username	description length	external URL	private	#posts	#followers	#follows	fake
1	0.27	0	0	0	53	0	0	32	1000	955	0
1	0	2	0	0	44	0	0	286	2740	533	0
1	0.1	2	0	0	0	0	1	13	159	98	0
1	0	1	0	0	82	0	0	679	414	651	0
1	0	2	0	0	0	0	1	6	151	126	0
1	0	4	0	0	81	1	0	344	669987	150	0
1	0	2	0	0	50	0	0	16	122	177	0
1	0	2	0	0	0	0	0	33	1078	76	0
1	0	0	0	0	71	0	0	72	1824	2713	0
1	0	2	0	0	40	1	0	213	12945	813	0

Fig -1: Training dataset

3.2 Artificial Neural Network

ANNs are computational models that are designed to emulate the human brain. The working of the brain is used as a basis by the ANN for development of complex algorithms that can be used to solve a given problem. Like the human brain ANN can decipher and arrive to a conclusion from a given vague information. ANN is made up of thousands of artificial neurons which are simply called units or nodes. These nodes are like the natural neurons both in structure and working [10].These nodes are connected to each other to form layers and the interconnection of these layers produces a web like structure which is called a network.

A fundamental ANN comprises of input, hidden and output layers. The input layer accepts input from the outside world. The nodes present in the input layer are passive in nature, this means that they are incapable of making any changes in the data provided to them. The input provided to this layer can be in form of a pattern or a vector in case of visual data. This accepted input is then transformed into something meaningful by the hidden layers. The hidden layers refine the features of the input before passing them on to the output layer. Finally the output layer responds to the given information and produces output for the system. Based on the number of hidden layers an ANN can either be shallow neural network or deep neural network. An ANN becomes shallow if it has a single hidden layer between the input and the output layer and it becomes deep it has at least 2 hidden layers in between the layers taking information and the layer giving yield. Since deep network has more than a single hidden layer learning becomes faster. Fig-2 depicts this difference. The number of hidden layers in an ANN model should be decided in such a way that it avoid both over-fitting and under-fitting. A large number of hidden layers mean that the model will learn all the details of the training set and so will perform very well on it but it will perform poorly on a data set which it has never seen before. This condition is called over-fitting. On the other hand too few or no hidden layers will make it difficult for the model to capture even the basic information about the training set. As a result the model will perform badly on both the training and testing set, resulting in the condition of under-fitting.

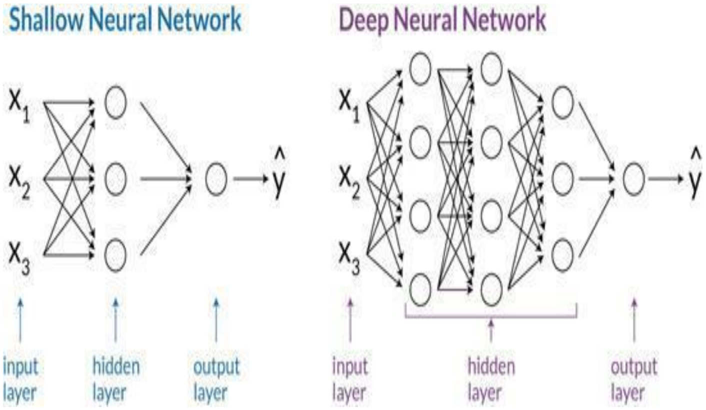


Fig -2: Shallow and deep neural network architecture

Apart from the number of hidden layers we can also use the direction of flow of information to classify an ANN. Based on this criteria an ANN can be feed forward or feed backward. In a feed forward network the information can only flow in forward direction, i.e., it is unidirectional and has no feed back loops. On the contrary feed back network allows the existence of feed back loops in its structure. Fig-3 shows feed forward and feed back network [17].

Each neuron is connected to all the neurons in the previous and the next layer. Thus the neurons are said to be interconnected to each other. Each interconnection between the neurons has weight associated to it. This associated weight can be positive, negative or even zero. The output of a neuron get multiplied to the weight associated with the interconnection through which it travels to the next neuron as input. Thus the strength of an information signal gets altered after it is multiplied with the weight. The signal might become weaker or stronger depending on the weight [11]. The strength of a signal might even become zero if the associated weight is zero. The model tries out a number of combinations of weights during the training process to find a combination which gives the most accurate prediction [12]. All the inputs arriving at a neuron are multiplied with their corresponding weights and then added. This weighted sum is then added to a bias to scale up the response of the system or to make the weighted sum non zero incase it takes value as zero. After addition of bias to the weighted sum it is passed through an activation function. An activation function is a transfer function which is used to get desired output from the weighted sum. It acts as a gate between the input of the current neuron and its output going to the next layer. It can be linear like sigmoid or non linear like softmax depending on the need. Fig-4 shows the working of a single node. In the figure shown 'b' is the bias, 'y' is the output, 'F' is the activation function used, 1 to xm are inputs and w0 to wm are the weights.

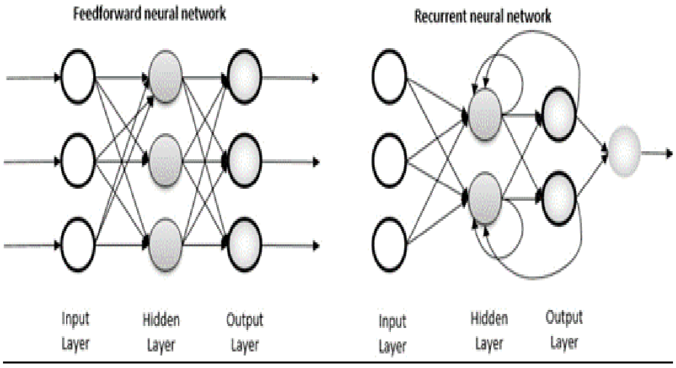


Fig -3: Feed forward and feed back network

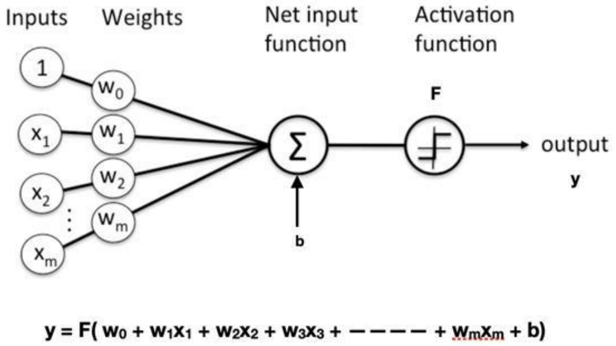


Fig -4: Working of a single node

An ANN has to go through a training phase before it can actually start making predictions. During the training ANN works on recognising the various patterns present in the data, these patterns can be visual, textual or even vocal.

Incase of unsupervised learning ANN simply learns to group the given data in clusters depending on the similarity shown by them. Incase of supervised learning the ANN model checks the output predicted during the training with the actual output while adjusting the weights to obtain the most accurate structure. Another type of learning is reinforcement learning where the model interacts with the environment to get its state and then choses an appropriate action to change its state. This chosen action gets sent to a simulator which gives feedback to the model in form of a numerical reward that can take positive or negative value. Based on the review received from the simulator the model changes or keeps its action. The model uses trail and error approach to find the most suitable action [18]

Once the model is trained it can be used to predict the result for an input set which it has never seen before.

ANN has been in existence since 1943 but it has come into limelight only in the recent years. The ability of ANN to go beyond the provided information in order to produce extremely accurate results is unmatched. Even data loss fails to affect the working of ANN as it already stores the data its network, thus eliminating the need of any database to keep a track of the data. These advantages of ANN have made it extremely popular among the researchers who have actively started exploring ANN for a number of application including image processing, text recognition, speech recognition, routing and translation.

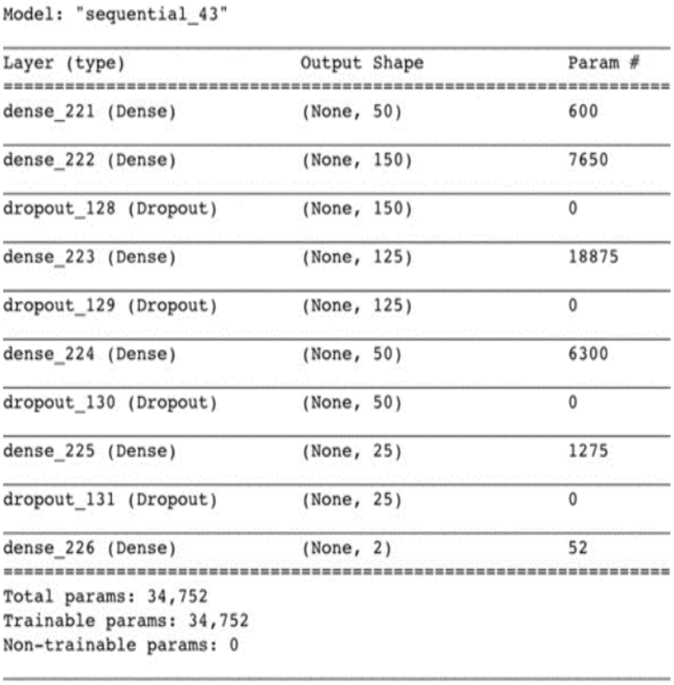
3.3 Model used

The ANN model used in this project is a 6 layered model. It has an input layer, 4 hidden layers and an output layer. The number of neurons in each layer has been set to reach the desired accuracy. The first layer which is the input layer has 50 neurons it is followed by the 1st hidden layer which has 150 neurons, it is further followed by the 2nd hidden layer which has 125 neurons, the next layer which is the 3rd hidden layer has 50 neurons, the 4th hidden layer is next having 25 neurons and the last layer which is the output layer has 2 neurons. The model takes 11 input variables. A dropout layer of 30 percent rate is applied after each hidden layer. All the hidden layers utilise Rectified Linear Unit (ReLu) function which is a linear function. It gives the input value provided to it as output if it is positive and gives 0 as the output if the value provided to it is negative or 0. The ReLu function is defined in Eq.1, shown as follows :-

$$y_i = \begin{cases} x_i & \text{if } x_i \geq 0 \\ 0 & \text{if } x_i < 0. \end{cases} \tag{1}$$

The output layer utilises softmax function which gives either 1 or 0 as output irrespective of the value of the input provided to it. It normalises the value of a K value into a K value vector having sum 1. The softmax function is defined in Eq.2 shown as follows :-

$$f_i(\vec{a}) = \frac{e^{a_i}}{\sum_k e^{a_k}} \tag{2}$$



```
Model: "sequential_43"
_________________________________________________________________
Layer (type)                 Output Shape              Param #
=================================================================
dense_221 (Dense)            (None, 50)                600
_________________________________________________________________
dense_222 (Dense)            (None, 150)               7650
_________________________________________________________________
dropout_128 (Dropout)        (None, 150)               0
_________________________________________________________________
dense_223 (Dense)            (None, 125)               18875
_________________________________________________________________
dropout_129 (Dropout)        (None, 125)               0
_________________________________________________________________
dense_224 (Dense)            (None, 50)                6300
_________________________________________________________________
dropout_130 (Dropout)        (None, 50)                0
_________________________________________________________________
dense_225 (Dense)            (None, 25)                1275
_________________________________________________________________
dropout_131 (Dropout)        (None, 25)                0
_________________________________________________________________
dense_226 (Dense)            (None, 2)                 52
=================================================================
Total params: 34,752
Trainable params: 34,752
Non-trainable params: 0
_________________________________________________________________
```

Fig -5: Model Summary

The model is trained using 20 epochs, this means that the whole dataset is passed forward and backward to the model 20 times in order to train it. Fig-5 shows the model summary.

4. RESULT

The proposed model gives an accuracy of 93.63 percent with 0.18 percent loss. To further study the performance of the model we can use confusion matrix and learning curves.

4.1 Confusion Matrix

Using confusion matrix we can describe the performance of the classification model [13]. Fig. 6 shows the confusion matrix for the deep learning ANN model proposed in this project. It can be further used to find out precision [14] and recall [14] of the model.

Precision is the ratio of true positives to the values predicted correctly. It is defined in Eq.3, given as follows :-

$$precision = \frac{true\ positives}{true\ positives + false\ positives} \tag{3}$$

Recall is the ratio of true positives to the total number of positives. It is defined in Eq.4, given as follows :-

$$recall = \frac{true\ positives}{true\ positives\ + false\ negatives} \tag{4}$$

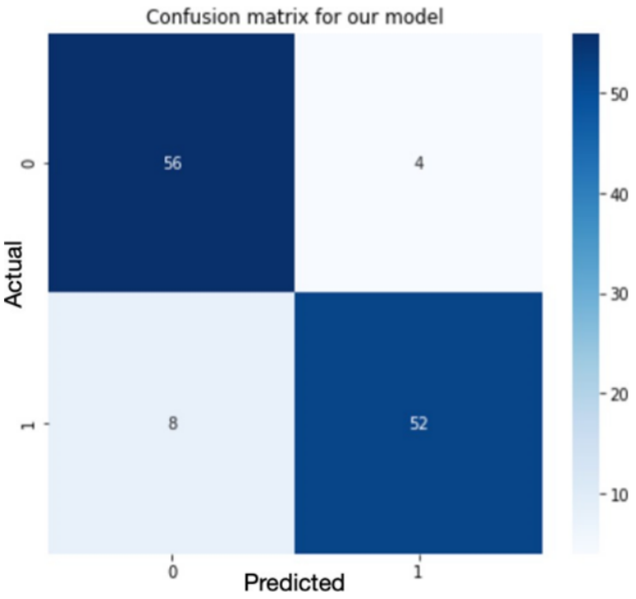


Fig -6: Confusion matrix for the model

4.2 Accuracy and loss of the model

Graphical outputs of the model can be used to check if it is over-fitted, under-fitted or perfectly fitted [15]. Fig-7 depicts the loss as the training process advances.

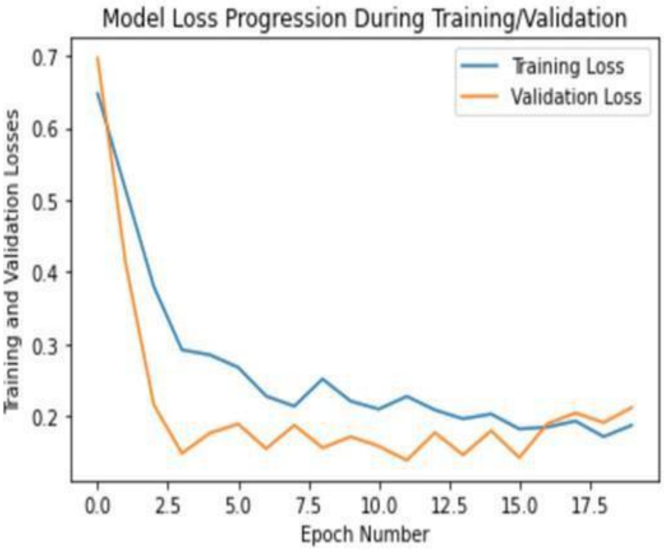


Fig -7: Graphical representation of losses wrt epochs

It can be observed that the training loss decrements with increment in epochs, same can be said for validation loss. Fig-8 represents the accuracies of the model which increments with epochs. In order for the model to be efficient the training and validation loss as well as accuracy should be comparable.

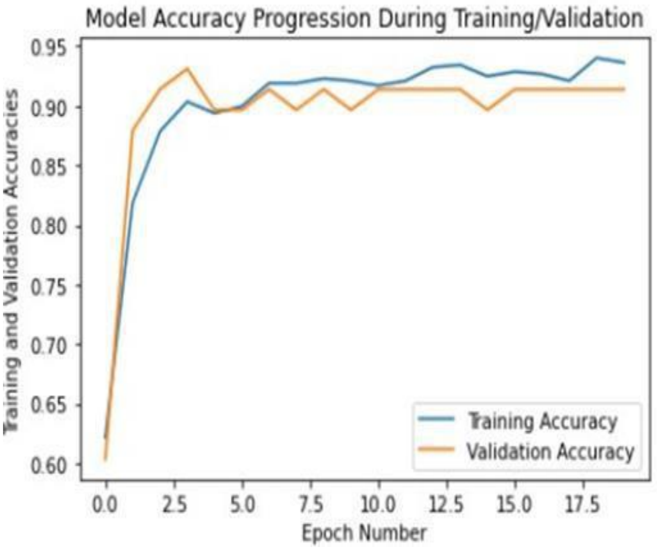


Fig -8: Graphical representation of accuracies wrt epochs

The training and validation data of the model lie in comparative ranges making it proficient [16].

5. CONCLUSIONS

In this research we have recognized a serious issue haunting the social media platforms which is the ever increasing number of fake accounts on them. To overcome this problem we have proposed a deep learning model which can be used to identify the dummy accounts in matter of seconds which can be then removed before they cause any serious harm to the people. The suggestion of a deep learning has been done in this project keeping in mind the drawbacks of the currently existing methods. The model used studies the data associated with the accounts to derive a relation between it and the genuineness of the account. To represent the performance of the model we have used confusion matrix. and learning curves along with the accuracy of the model. The model has shown good performance incase of both training and testing set.

Currently only the data available for Instagram profiles has been used for the training and testing purpose but in future we can also train the model to identify fake accounts on other popular platforms like Facebook, LinkedIn, Twitter and many more by providing an efficient dataset for them.

REFERENCES

[1] Edosomwan, Simeon & Prakasan, S.K. & Kouame, D. & Watson, J. & Seymour, T.. (2011). The history of social media and its impact on business. Journal of Applied Management and Entrepreneurship. 16. 79-91.

[2] Suja P Mathews,Sunu George,"Growth and Future of Social Media",International Journal of Advanced Research in Computer Engineering & Technology (IJARCET) Volume 2, Issue 12, December 2013.R. Nicole, "Title of paper with only first word capitalized," J. Name Stand. Abbrev., in press.

[3] Adikari, Shalinda, and Kaushik Dutta. "Identifying Fake Profiles in LinkedIn." In PACIS, p. 278. 2014.

[4] Bharat Sampatrao Borkar, Dr. Rajesh Purohit,"Recognition of fake profiles in social media : a literature review",Volume 5, Issue 2, 2019

[5] Raturi, Rohit. (2018). Machine Learning Implementation for Identifying Fake Accounts in Social Network.

[6] Samala Durga Prasad Reddy,"Fake Profile Identification using Machine Learning" in IRJET 2019

[7] 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.

[8] S. P. Maniraj, Harie Krishnan G, Surya T, Pranav R, "Fake Account Detection using Machine Learning and Data Science", International Journal of Innovative Technology and Exploring Engineering (IJITEE) ISSN: 2278-3075, Volume-9 Issue-1, 2019

[9] Ananya Dey , Hamsashree Reddy, Manjistha Dey and Niharika Sinha, "Detection of fake accounts in Intagram using machine learning", International Journal of Computer Science & Information Technology (IJCSIT) Vol 11, No 5, October 2019.

[10] Shiruru, Kuldeep. (2016). An Introduction To Artificial Neural Network. International Journal of Advance Research and Innovative Ideas in Education. 1. 27-30.

[11] B. Arora, et.al, "Application of Artificial Neural Network in Cryptography", International Conference on " Power Energy, Environment and Intelligent Control(PEEIC), October 18-19, 2019

[12] Agatonovic-Kustrin, S., & Beresford, R. (2000). Basic concepts of artificial neural network (ANN) modeling and its application in pharmaceutical research. Journal of Pharmaceutical and Biomedical Analysis , 22 (5) , 717–727.https://doi.org/10.1016/s0731-7085(99)00272-1

[13] Visa, et al. : "Confusion Matrix-based Feature Selection." MAICS 710, 120-127, 2011.

[14] Davis, J., and Goodrich, M. : The relationship between Precision- Recall and ROC curves. 23rd international conference on machine learning, 233-240, 2006

[15] JabbarH. And Khan R.Z., : Methods to avoid over-fitting and under- fitting in supervised machine learning (comparative study). Computer Science, Communication and Instrumentation Devices, 163-172, 2015

[16] Michel Jose Anzanello, Flavio Sanson Fogliatto, Learning curve models and applications: Literature review and research directions,International Journal of Industrial Ergonomics,Volume 41, Issue 5,2011

[17] Pekel, Engin & Kara, Selin. (2017). A Comprehensive Review For Artifical Neural Network Application To Public Transportation. Sigma Journal of Engineering and Natural Sciences. 35. 157-179.

[18] Fabiyi, Samson Damilola. (2019). A Review of Unsupervised Artificial Neural Networks with Applications. International Journal of Computer Applications. 181. 22-26. 10.5120/ijca2019918425.