International Journal For Advanced Research In Science & Technology
A peer reviewed international journal www.ijarst.in
ISSN: 2457-0362

IDENTIFING OF FAKE PROFILES ACROSS ONLINE SOCIAL NETWORKS BY USING NEURAL NETWORK

CH.VENKATESH [1],P.VEDA SRI[2], R.VAISHNAVI[3], A.SOWJANYA[4], M.NAVYA SREE[5] , MD.IBRAHIM UL HASSAN [6]

[1] *Assistant Professor*,DepartmentofCSE,***ARJUN COLLEGE OF TECHNOLOGY AND SCIENCE***,Batasingaram(vill),hayathnagar(MDL),Rangareddy(dist)Telangana India. 501512

[2, 3, 4,5,6]UG Student,DepartmentofCSE,***ARJUN COLLEGE OF TECHNOLOGY AND SCIENCE***,Batasingaram(vill),hayathnagar(MDL),Rangareddy(dist)Telangana India. 501512

Abstract:

In seeing the present condition, online social networks are engaging with the majority of the people. From child to adult, all are spending a considerable time on these platforms either by exchanging information or making efficient communication with others. But nowadays, these social networking sites are suffering from a lot of fake accounts in taking advantage of vulnerabilities, either taking the benefits or targeting accounts attempting cybercrimes.

I. INTRODUCTION:

1.1 Research Problem

The concern about fake profile is protecting personal data or information from cyber attacks known as phishing attacks. The cyber attackers are often use this in stealing of information. In detecting of passwords, sharing of irrelevant contents, raising awareness this type of profiles are involved in all unlawful activity. In managing and taking the advantages of the critical situation this can be lead to the anonymity through a longer way. For reducing the incidents like trolling, hacking and cyber bullying this is need to be identified.

1.2 Research Rationale

In securing the all types of social accounts and keeping the users away from the cyber hackers this is necessary to identify those and using of ANNs model this can improved in more better way.

1.3 Research Approach

In regards to this, an "artificial neural network" system has been introduced as a part of the computer system. It is designed for simulating in a way in which the human brain possesses and analyses information. The inductive research approach can be considered for this type. In viewing the existing process and situations this can be observed through the patterns and system regularities. In taking the technical advantage ANN model need to be used effectively. It can be described as a foundation of artificial intelligence which will solve the problem in proving the difficulty according to human standards. Therefore "artificial neural networks" (ANNs) are introduced as a process of modeling, allowing the human nervous system through learning technique. By depending on the prediction, this detection process is revealing about the "user-level activities' '. User influence is also vital in reporting about the abnormalities. The social influence upon users can be assessed with the two types of factors. One is to find the user's impact upon others, and the other is to give the user importance. The evaluation is also based on the "fine-grained feature'.

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
www.ijarst.in
ISSN: 2457-0362

1.4 Aim and Objectives

The main aim is defined in reducing of fake accounts in social network by involving of ANN process. The objectives are –

- To identify of fake accounts in social networks
- To apply the ANNs through machine learning technique

Figure 1: Social Network

(Source:http://www.webtrafficroi.com/wp-content/uploads/2012/06/social-networks.jpg)

II. Introduction:

Literature Review

In viewing Ramalingam and Chinnaiah (2017), most social networking sites cannot notify the fake profiles. Therefore the discrimination can be seen in between the fake and real profiles, which is technically challenging for most of the users. The existing model is used for this related research study. From the available dataset, each of the social sites is extracting features by using the component analysis. Apart from this, the"Sybil frame" can be used as a multi-level mechanism which is detecting the Sybil's of Twitter and Facebook (Hajdu*et al.* 2019). There are also two types of approaches available for this type: structure-based and content-based. The "vote trust" is available for the identification of the mechanism for the classification between fake and benign user accounts

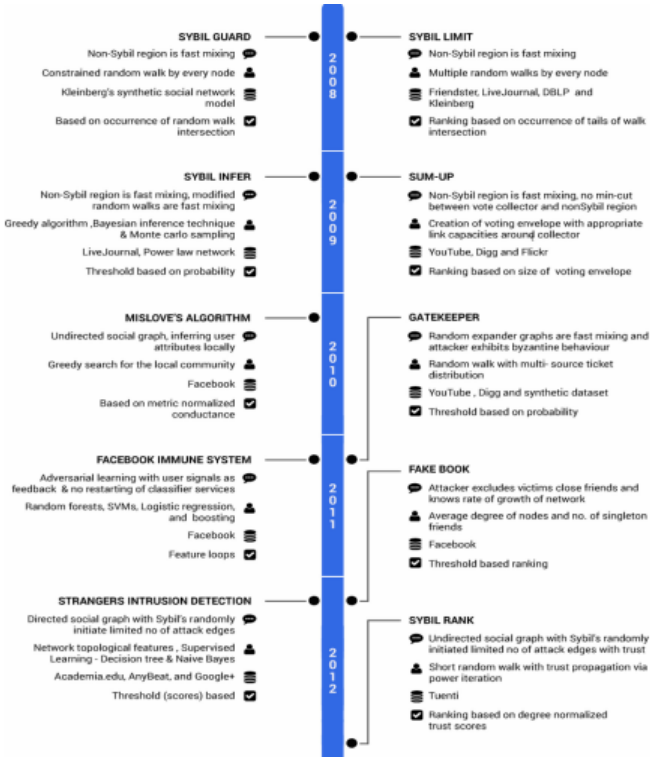


Figure2: Fake Profile Detection Models

(Source:https://e-tarjome.com/storage/panel/fileuploads/2019-06-13/1560405854_E11304-e-tarjome.pdf)

The rank algorithm can be used for the analysis of the user's influence upon his friends. For the analysis of data from the various types of OSNs, big data can be introduced. By finding the challenges, the capability of algorithms is finding through the reduction of computational cost, time complexity, development of performance and enabling the local learning techniques. Recent experiments are also revealing the related difficulties with the "in-memory management" of big data (Ramalingam

and Chinnaiah, 2018). This is included with the "Hekton", "SAP HANA", 'H-Store" and more. The "in-memory" data processing is also included with the big data analytics such as Spark, "Main Memory Map Reduce (M3R)" and real-time processing systems such as "Yahoo Simple Scalable Streaming System". In finding the challenges about data management, this can be found by finding the indexing possibilities, controlling concurrency, overflowing of data, and query processing.

According to Saatvik*et al.* 2020, the artificial neural network is mainly depending upon the three main factors, which are organizing the structure, institution and data component of a unit and data affiliation largeness. If one of these two parameters is right, then ANN lead can be described with the heap estimation. From the outset to self-assertive characteristics, the "heaps of net" are readied. The characteristics of a defined case's commitment can be easily determined by the data units (Khaled *et al.*, 2018). The yield of the net can be adjusted through a certain degree in a way through which the net is adjusted with a similar degree type. By bringing the yield estimation into the bet closer, the perfect yield's characteristics can be determined. In relating to this, the neutral structure of action can be introduced with consolidates with breaking down of the picture, "convolution neural structure", and a guide "neutral system. Therefore the organizing map can be provided with the quantization of the picture test. Within the topological position, the input is closed to the head; spaces are shut in the yield space. The best mirror of improved acknowledgement and 3D face geometry is related to the "Spectral Regression Kernel DiscriminateAnalysis" (SRKDA). This SRKDA is also subjected to a loss of faith where the apparition of diagram assessment can be presented with the proposed method (Awasthi*et al.* 2020). The careful approaches cannot loosen up with the minimal and dimensional size issues, which will further improve the fuse extraction through the non-direct structures. Artificial intelligence is built on ANNs through which are disrupting the multiple traditional ways of doing things. By using learning algorithms, the ANN can incorporate different sectors.

III. IMPLEMENTATION:

MODULES:

Module Details:

1.Upload Social Network Profiles Dataset:

Using this module we will upload dataset to application

2.Preprocess 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

3.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.

4.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.

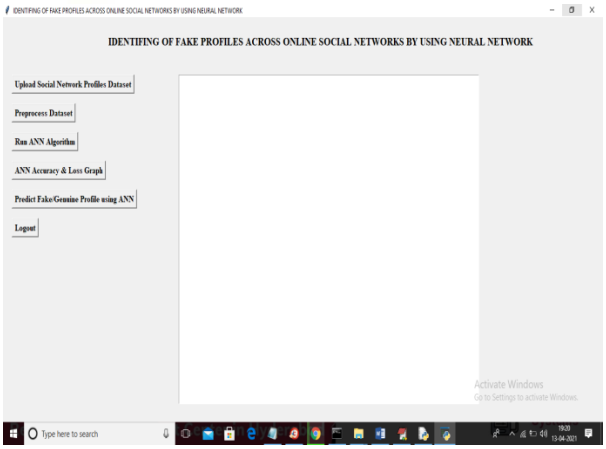
5.Predict Fake/Genuine Profile using ANN:

using this module we will upload new test data and then apply ANN train model to predict whether test data is genuine or fake.

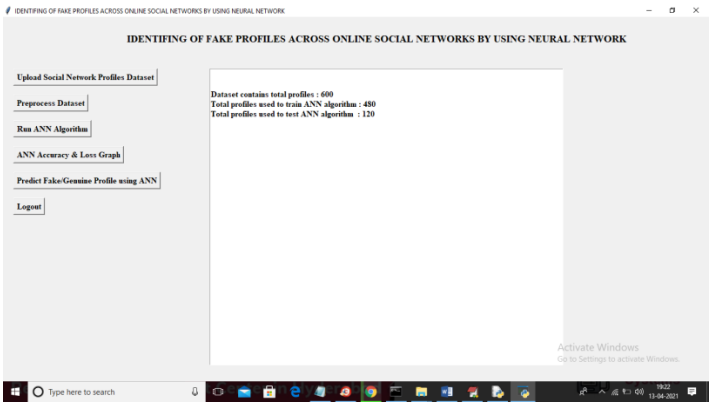
IV.SCREENSHOTS

To run project double click on 'run.bat' file to get below screen

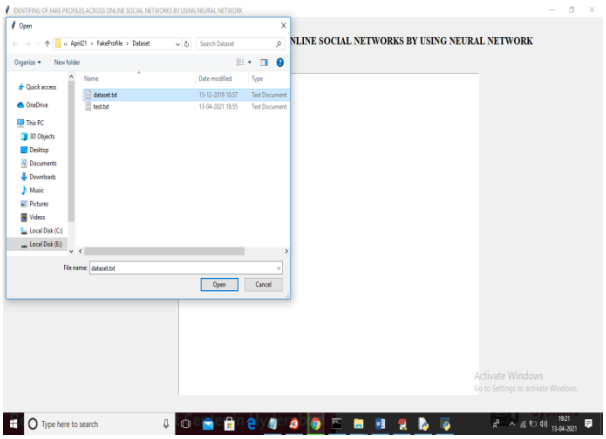
International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
www.ijarst.in
ISSN: 2457-0362



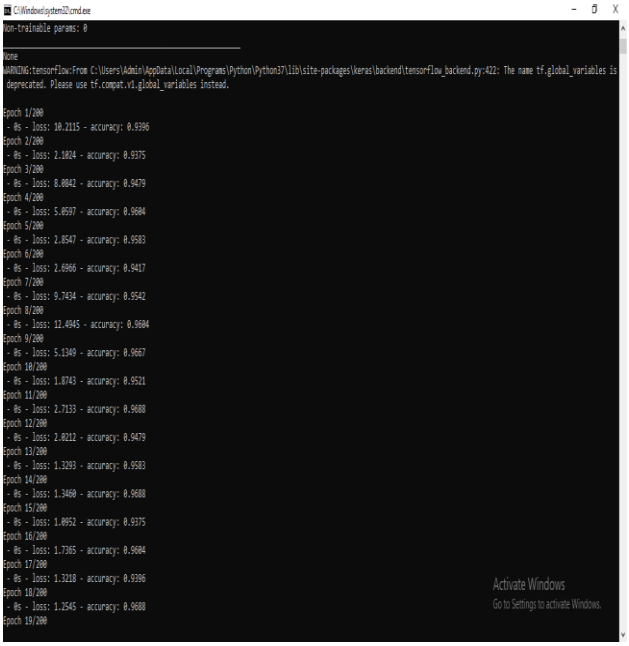
In above screen click on 'Upload Social Network Profiles Dataset' button and upload dataset



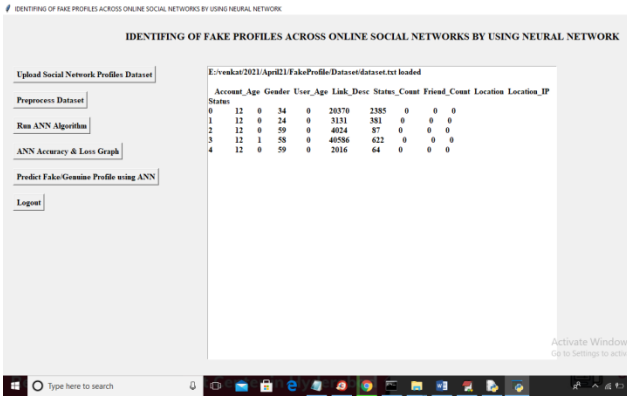
In above screen we can see dataset contains total 600 records and application using 480 records for training and 120 records to test ANN and now dataset is ready and now click on 'Run ANN Algorithm' button to ANN algorithm



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

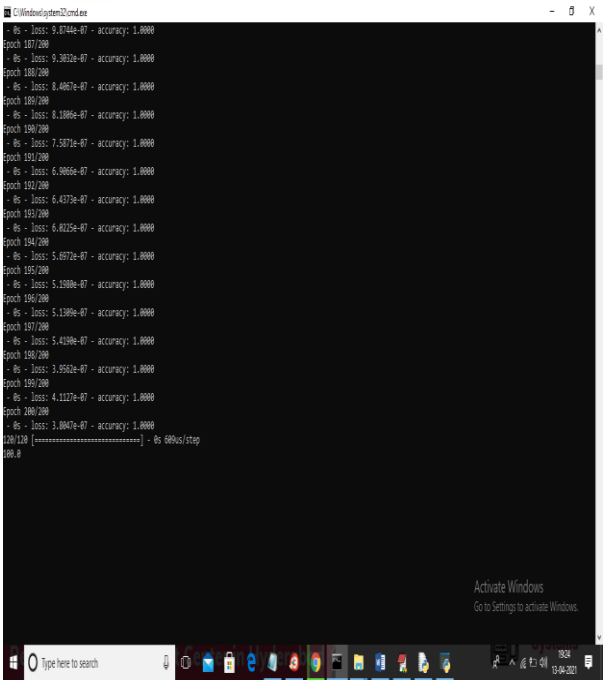


In above screen we can see ANN start iterating model generation and at each increasing epoch we can see accuracy is getting increase and loss getting decrease.

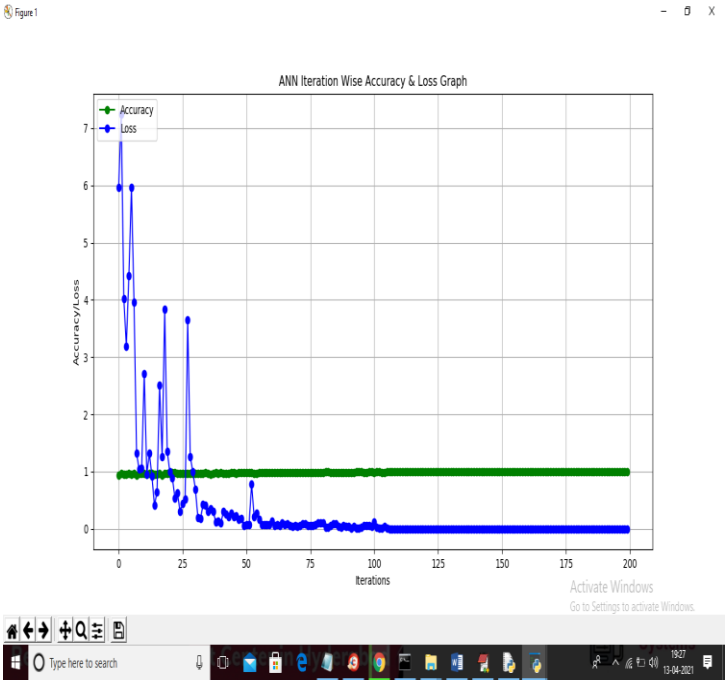


In above screen dataset loaded and displaying few records from dataset and now click on 'Preprocess Dataset' button to remove missing values and to split dataset into train and test part

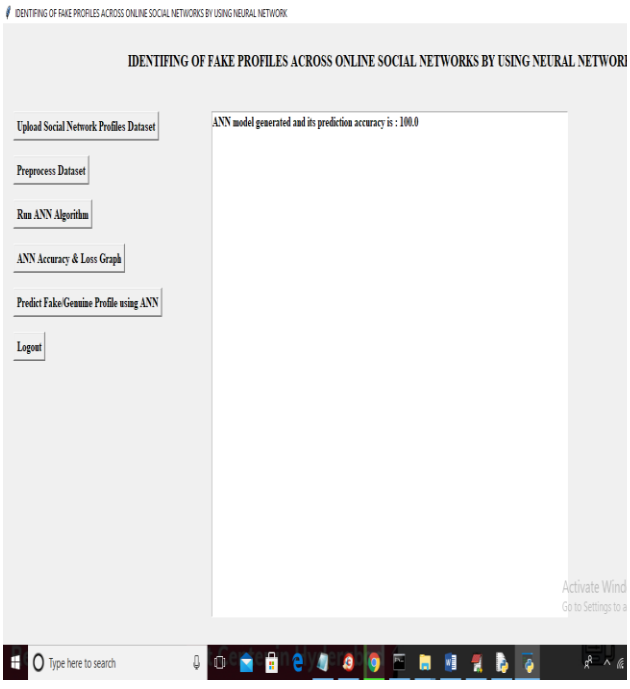
International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
www.ijarst.in
IJARST
ISSN: 2457-0362



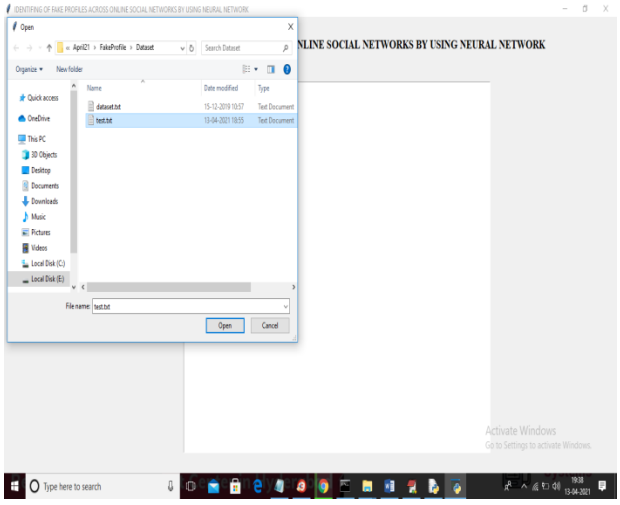
In above screen we can see after 200 epoch ANN got 100% accuracy and in below screen we can see final ANN accuracy



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

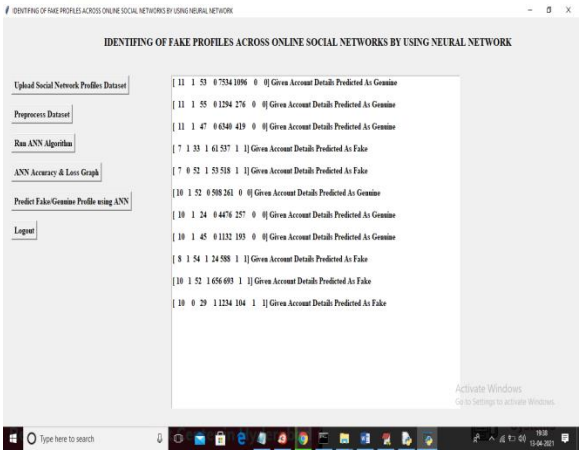


In above screen ANN model generated and now click on 'ANN Accuracy & Loss Graph' button to get below graph



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

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
www.ijarst.in
ISSN: 2457-0362



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

Significance of Research

ANNs are basically known as lone performers, which are not intended in the production of the general network types. This software is used for practical application through its networks. The primary focus is on forecasting and data mining. The software tools are used as -

1. Darknet,
2. NeuroSolutions,
3. Neural Designer,
4. Keras,
5. Neuroph,
6. Tflearn,
7. Torch,
8. “Stuttgart Neural Network Simulator”,
9. ConvNetJS,
10. NVIDIA DIGITS,

The ANN process has the ability in the relearning process according to the newer data types. Due to the uncertainty and complexity, it is difficult in defining a particular analytical model. In the elaborate ritual, a powerful computer-based application can be used. Therefore, the optimization technique's principle lies in the optimization process through which both the constraints and object functions are evaluated into the simulation model (Wanda and Jie, 2020). For the combined simulation, the optimization techniques and ANN need to be provided with the practical means for higher complex optimization. In searching for the solution space, the 'multi-objective optimization algorithm" or NSGA-II is used with adaptive local search. In discrediting the event simulation model, both input-output data is used for the generation of ANN in approximating the object function. Acting as an intelligent brain, this can train simulated data and accurate models.

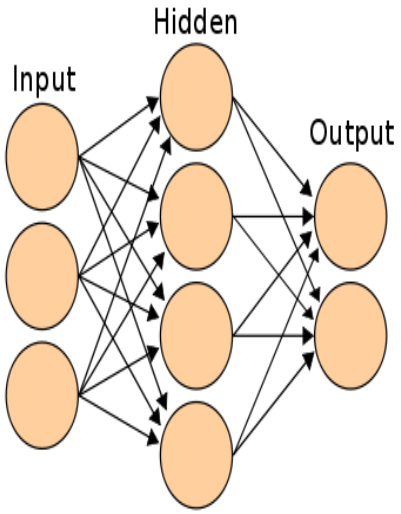


Figure3: ANN Framework

(Source:https://www.analyticsvidhya.com/wp-content/uploads/2014/10/ANN.png)

In between the nodes, the linkages are considered as the main factors (Zhang *et al.* 2020). By finding random weights of the linkages at the start of the algorithm, using the inputs for finding the linkages, searching the errors at the output nodes, weight calibration in

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
www.ijarst.in
ISSN: 2457-0362

between hidden and input nodes, defining the final linkage weights for the scoring of activation rate the framework is structured. Apart from this, by using hidden nodes and their linkages with the output, the output nodes' activation nodes can be found out.

4: Required Resources

In finding the resources, there are multiple modules that can be used. As the social network is a general site, by implementing artificial neural networks, different kinds of modules can be used in the detection of fake profiles. PyBrain is known as a modular within the machine learning library in using Python. Comparing the algorithm with predefined environments can offer better machine learning tasks. Scikit-learn are used for machine learning through Python (Meligy*et al.*, 2017). In predictive data analysis, it is considered as efficient tools. The sexmachine was created for publishing Python 3 compatible versions into PyPi. Without bugging, it can add definite improvements. In relation to this matplotlib is considered a comprehensive library used for animated, static, and interactive visualizations in python. This can make easy and more challenging things more efficiently to create. The ipython notebook is also known as the Jupiter notebook. In the computational environment, it can be combined with the execution of codes, plots and mathematics. Therefore ipython is also known as an interactive shell of python. A Jupiter kernel works with the code in the notebook.

Section 5: Required Skills

The activity of this related technique is from translating web pages into three virtual assistants to order groceries while conversing with chatbots in solving problems. Email servers are also using ANNs and deleting spam from the user inbox. Chatbots are also developed with ANNs as a "natural language processing". Pandas in the package of python are delivering their flexible, fast and flexible data structure in working with the level data types. For working with the array types, NumPy is used in the python library. The working function lies in the domain of "linear algebra". In relation to this pipe is known as the package manager of python. This is used as a distributed part of the standard library (Kaur and Sabharwal, 2018). Apart from this, the knowledge is required about Java, Python more clearly. In using the modules and packages, depth knowledge and preferred system configuration are preferable. Python 3.9 is used as the best version. Therefore, the ram, hard disk capacity, and the IDE packages are necessary for working with the programs.

6: Project Plan

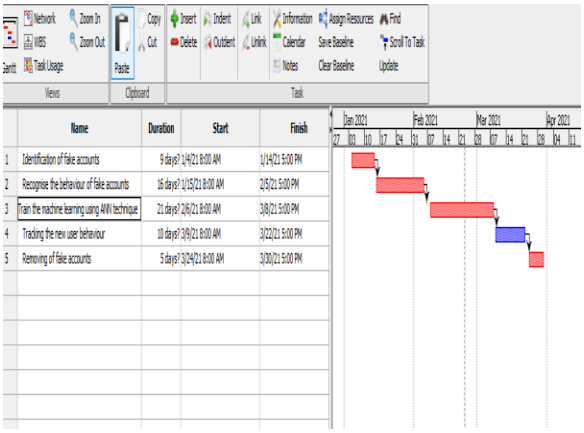


Figure4: Gantt chart

(Source: Project Libra)

ReferenceJournals

Awasthi, S., Shanmugam, R., Jena, S.R. and Srivastava, A., 2020. Review of Techniques to Prevent Fake Accounts on Social Media.

Hajdu, G., Minoso, Y., Lopez, R., Acosta, M. and Elleithy, A., 2019, May. Use of Artificial Neural Networks to Identify Fake Profiles. In *2019 IEEE Long Island Systems, Applications*

and Technology Conference (LISAT) (pp. 1-4). IEEE.

Kaur, J. and Sabharwal, M., 2018. Spam detection in online social networks using feed forward neural network. In *RSRI conference on recent trends in science and engineering* (Vol. 2, pp. 69-78).

Khaled, S., El-Tazi, N. and Mokhtar, H.M., 2018, December. Detecting fake accounts on social media. In *2018 IEEE International Conference on Big Data (Big Data)* (pp. 3672-3681). IEEE.

Meligy, A.M., Ibrahim, H.M. and Torky, M.F., 2017. Identity verification mechanism for detecting fake profiles in online social networks. *Int. J. Comput. Netw. Inf. Secur.(IJCNIS)*, *9*(1), pp.31-39.

Ramalingam, D. and Chinnaiah, V., 2018. Fake profile detection techniques in large-scale online social networks: A comprehensive review. *Computers & Electrical Engineering*, *65*, pp.165-177.

Wanda, P. and Jie, H.J., 2020. DeepProfile: Finding fake profile in online social network using dynamic CNN. *Journal of Information Security and Applications*, *52*, p.102465.

Zhang, J., Dong, B. and Philip, S.Y., 2020, April. Fakedetector: Effective fake news detection with deep diffusive neural network. In *2020 IEEE 36th International Conference on Data Engineering (ICDE)* (pp. 1826-1829). IEEE.