JES
Journal of Engineering Sciences

Vol 13, Issue 02, FEB /2022

ISSN NO:0377-9254

Identification Of Fake Profiles in Online Social Networks Using Artificial Neural Network

Rajasekhar Nennuri[1], Nisha Kumari[2], Nikitha Reddy Donti[3], Naga Bharan Kaza[4]

#1 Assistant Professor,Dept. of Computer Science Engineering ,INSTITUTE OF AERONAUTICAL ENGINEERING, Dundigal, Hyderabad - 500 043, Telangana, India.

#2,#3,#4 Studenht ,Dept. of Computer Science Engineering ,INSTITUTE OF AERONAUTICAL ENGINEERING, Dundigal, Hyderabad - 500 043, Telangana, India.

Abstract_ These days, there is a noticeable increase in applied sciences. Mobile phones are becoming smarter. Technology is associated with online social networks, which have emerged as a part of everyone's life in terms of making new friends and retaining friends, as well as making their hobbies easier. However, the increase in online networking causes various problems, such as fabricating their profile. — In this research, we employ computing device learning, specifically a synthetic neural community, to identify whether or not a Facebook buddy request is legitimate. We also outline the training and libraries that will be used. We also discuss the sigmoid feature and how weights are determined and used.Finally, we reflect onconsideration on the parameters of the social community web page which are utmost necessary in the furnished solution.

1.INTRODUCTION

Long range interpersonal communication has end up a notable diversion inside the web as of now, drawing in countless clients, burning through billions of minutes on such administrations. Online Social organization (OSN) administrations assortment from social cooperations based stages like Facebook or MySpace, to understanding spread driven stages suggestive of twitter or Google Buzz, to Social association trademark brought to introduce frameworks like Flicker. The contrary hand, improving security concerns and safeguarding the OSN privateness actually connote a most significant bottleneck and saw mission. While utilizing Social organization's (Sn's), exceptional people share stand-out amounts of their private arrangement. Having our singular expertise totally or to some degree uncovered to the overall population, makes us amazing focuses for exceptional kinds of attacks, the most exceedingly terrible of which could be ID burglary. Data

fraud happens when any singular uses character's skill for a private achieve or reason. During the prior years, online recognizable proof burglary has been an essential issue thinking of it as impacted huge number of individuals' around the world. Casualties of ID robbery might experience remarkable sorts of punishments; for delineation, they would lose time/cash, get dispatched to reformatory, get their public picture destroyed, or have their associations with partners and friends and family harmed. As of now, by far most of SN's does no longer checks common users‟ obligations and has entirely vulnerable privateness and security approaches. Truth be told, most SN's applications default their settings to negligible privateness; and subsequently, SN's turned into a best stage for misrepresentation and misuse. Person to person communication contributions have worked with data fraud and Impersonation assaults for genuine comparable to innocent assailants. To compound the situation, clients are expected to outfit right comprehension to set up a record in Social Networking sites. Simple observing of what clients share on-line would prompt devastating misfortunes, not to mention, assuming such bills had been hacked. Profile data in web-based organizations will likewise be static or dynamic. The subtleties which can be provided with the guide of the individual on the hour of profile creation is known as static information, the spot as the important part that are described with the guide of the framework inside the organization is called dynamic information. Static information incorporates segment components of an individual and his/her advantages and dynamic information remembers individual runtime propensities and region for the organization. By far most of momentum research relies upon static and dynamic information. Anyway this isn't applicable to heaps of the informal organizations, where handiest some of static profiles are seen and dynamic profiles as a rule are not clear to the individual organization. In excess of a couple of methods have been proposed by exceptional specialist to understand the phony characters and malignant substance material in web-based informal communities. Each interaction had its own merits and negative marks. The issues including long range interpersonal communication like security, web based tormenting, abuse, and savaging and numerous others.

Are large numbers of the cases used by misleading profiles on long range interpersonal communication locales. Misleading profiles are the profiles which are not explicit i.E. They're the profiles of people with misleading qualifications. The misleading Facebook profiles all the more generally are enjoyed pernicious and unwanted exercises, bringing on some issues to the social local area clients. People make counterfeit profiles for social designing, online pantomime to stigmatize a man or lady, advancing and lobbying for a person or a horde of people. Facebook has its own security framework to monitor individual certifications from spamming, phishing, etc. Furthermore the equivalent is regularly called Facebook Immune framework (FIS). The FIS has now not been prepared to notice counterfeit profiles made on Facebook through clients to a greater degree.

2.LITERAURE SURVEY

Various fake record recognition methodologies depend on the investigation of individual interpersonal organization profiles, with the point of distinguishing the qualities or a combination thereof that help in recognizing the legitimate and the fake records. In particular, various features are extracted from the profiles and posts, and after that Machine learning algorithms are used so as to construct a classifier equipped for recognizing fake records. For instance,

Nazir et al. (2010) [1] describes recognizing and describing phantom profiles in online social gaming applications. The article analyses a Facebook application, the online game “Fighters club”, known to provide incentives and gaming advantage to those users who invite their peers into the game. The authors contend that by giving such impetuses the game motivates its players to make fake profiles. By presenting those fake profiles into the game, the user would increase a motivating force of an incentive for him/herself.

Adikari and Dutta (2014) [2] depict recognizable proof of fake profiles on LinkedIn. The paper demonstrates that fake profiles can be recognized with 84% exactness and 2.44% false negative, utilizing constrained profile information as input. Techniques, for example, neural networks, SVMs, and Principal component analysis are applied. Among others, highlights, for example, the number of languages

spoken, training, abilities, suggestions, interests, and awards are utilized. Qualities of profiles, known to be fake, posted on uncommon sites are utilized as a ground truth.

Chu et al. (2010) [3] go for separating Twitter accounts operated by humans, bots, or cyborgs (i.e., bots and people working in concert). As a part of the detection problem formulation, the Identification of spamming records is acknowledged with the assistance of an Orthogonal Sparse Bigram (OSB) text classifier that uses pairs of words as features.

Stringhini et al. (2013) [4] analyze Twitter supporter markets. They describe the qualities of Twitter devotee advertises and group the clients of the business sectors. The authors argue that there are two major kinds of accounts who pursue the "client": fake accounts("sybils"), and compromised accounts, proprietors of which don't presume that their followers rundown is expanding. Clients of adherent markets might be famous people or legislators, meaning to give the appearance of having a bigger fan base, or might be cybercriminals, going for making their record look progressively authentic, so they can rapidly spread malware what's more, spam.

Thomas et al. (2013) [5] examine black market accounts utilized for distributing Twitter spam.

De Cristofaro et al. (2014) [6] investigate Facebook like cultivates by conveying honeypot pages.

Viswanath et al. (2014) [7] identify black market Facebook records based on the examination of anomalies in their like behavior. Farooqi et al. (2015) [6] explore two black hat online commercial centers, SEO Clerks and My Cheap Jobs. Fayazi et al. (2015) think about manipulation in the online review.

3.PROPOSED SYSTEM

In our solution, we use machine learning, namely an artificial neural network to determine what are the chances that a friend request is authentic or not. We utilize Microsoft Excel to store old and new fake data profiles. The algorithm then stores the data in a data frame. This collection of data will be divided into a training set and a testing set. We would need a data set from the social media sites to train our model.

For the training set, the features that we use to determine a fake profile are Account age, Gender, User age, Link in the description, Number of messages sent out, Number of friend requests sent out, Entered location, Location by IP, Fake or Not. Each of these parameters is tested and assigned a value. For example, for the gender parameter if the profile can be determined to be a female or male a value of (1) is assigned to the training set for Gender. The same process is applied to other parameters. We also use the country of origin as a factor We then determine the Number of messages sent out parameter by dividing the number of messages sent by the age of the account. We then determine the Number of friend requests sent out parameter by dividing the Number of friend computing and used primarily for multi-dimensional matrix multiplication as we are dealing with a large amount of numbers that are very dependent on each other.

In this paper using Artificial Neural Networks we are identifying whether given account details are from genuine or fake users. ANN algorithm will be trained with all previous users fake and genuine account data and then whenever we gave new test data then that ANN train model will be applied on new test data to identify whether given new account details are from genuine or fake users.

Online social networks such as Facebook or Twitter contains users details and some malicious users will hack social network database to steal or breach users information, To protect users data we are using ANN Algorithm.

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

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

All fake users 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.

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 less number of posts as their main intention is to send friend requests not posts, so by analysing this features Facebook mark that record with value 1 which means it's a fake account. We are using above dataset to train ANN model and this dataset saved inside code 'dataset' folder. After building train model we input test data with account details and ANN will give result as fake or genuine. Below are some values from test data

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 there and ANN will predict status and give us result whether above test data is fake or genuine. In output we can see result of above test data.

3.1 IMPLEMENTATIONS

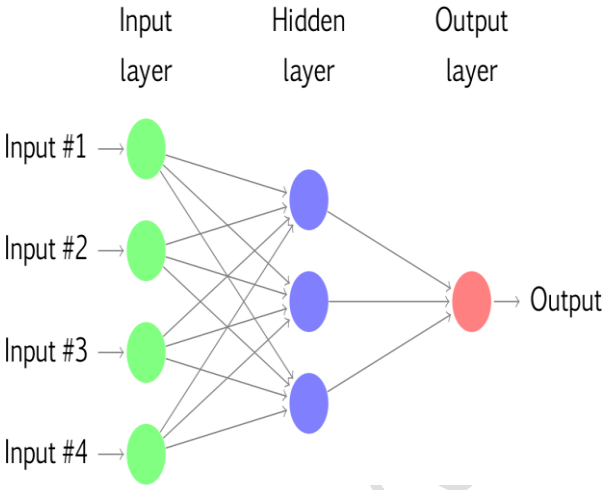
ANN algorithms Details

To demonstrate how to build a ANN neural network based image classifier, we shall build a 6 layer neural network that will identify and separate one image 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 image classification 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 convolutional 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 function 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



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

3.2 MODULE DETAILS:

Admin Module: Admin will login to application by using username as 'admin' and password as 'admin' and then perform below actions.

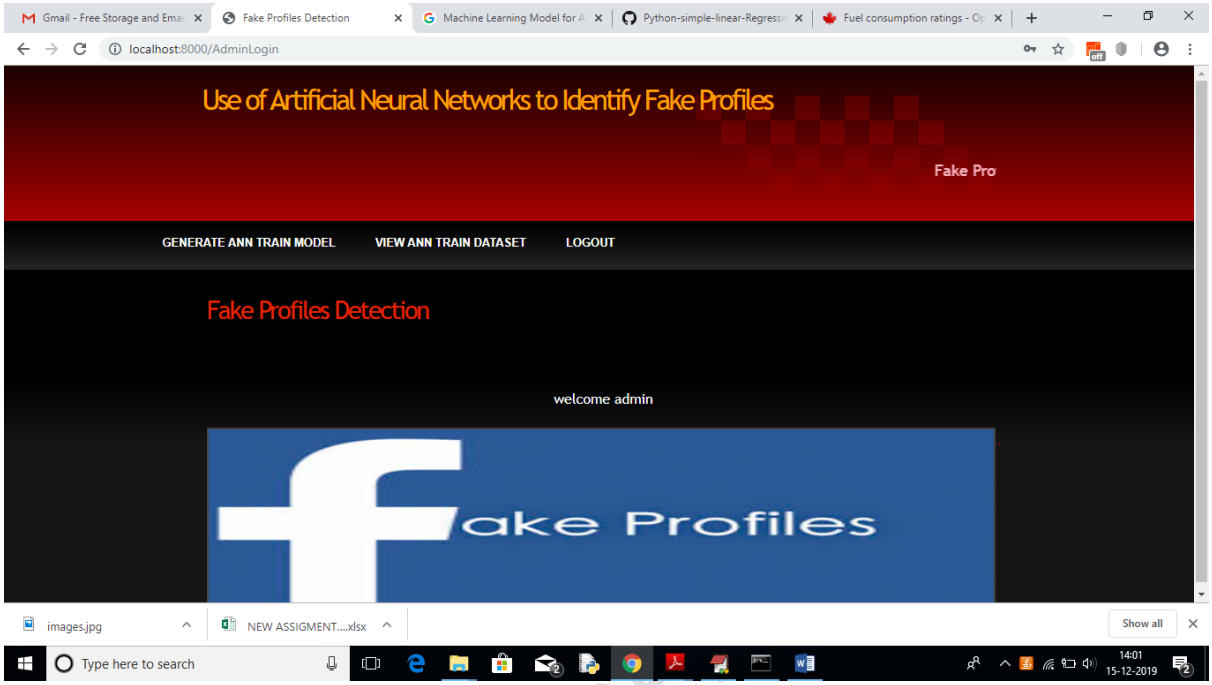
a) Generate ANN Train Model: Admin will upload profile dataset to ANN algorithm to build train model. This train model can be used to predict fake or genuine account by taking new account test data.

b) View ANN Train Dataset: Using this module admin can view all dataset used to train ANN model.

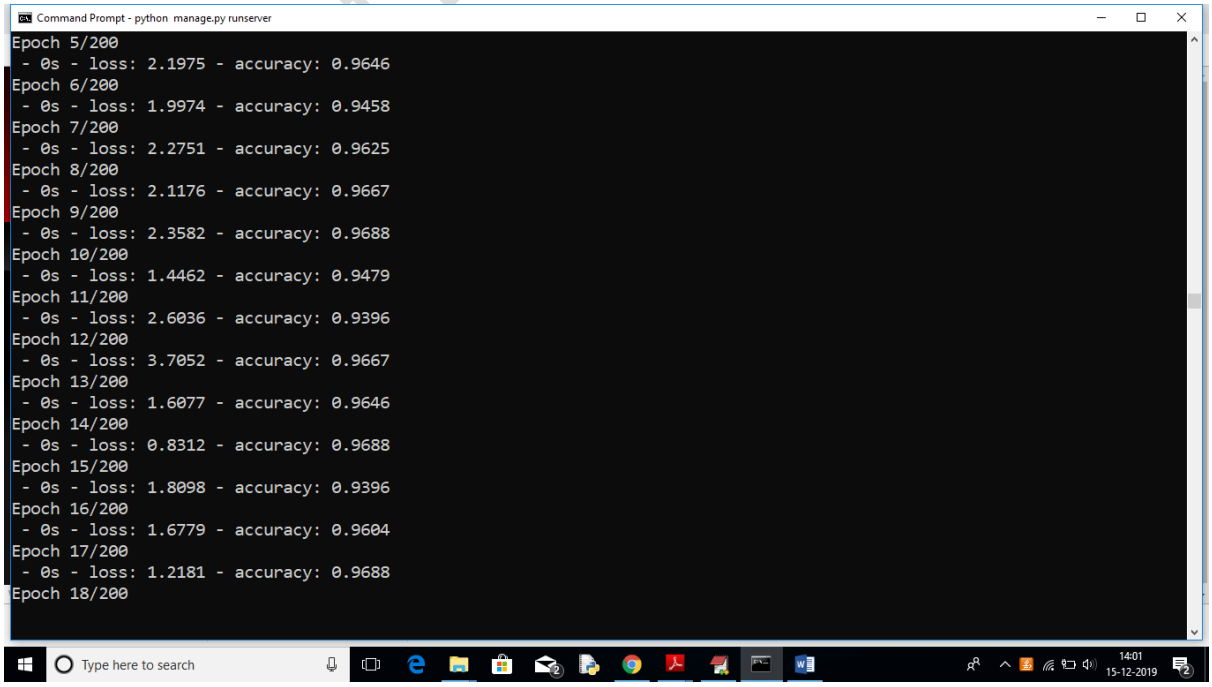
User Module: Any user can use this application and enter test data of new account and call ANN algorithm. ANN

algorithm will take new test data and applied train model to predict whether given test data contains fake or genuine details.

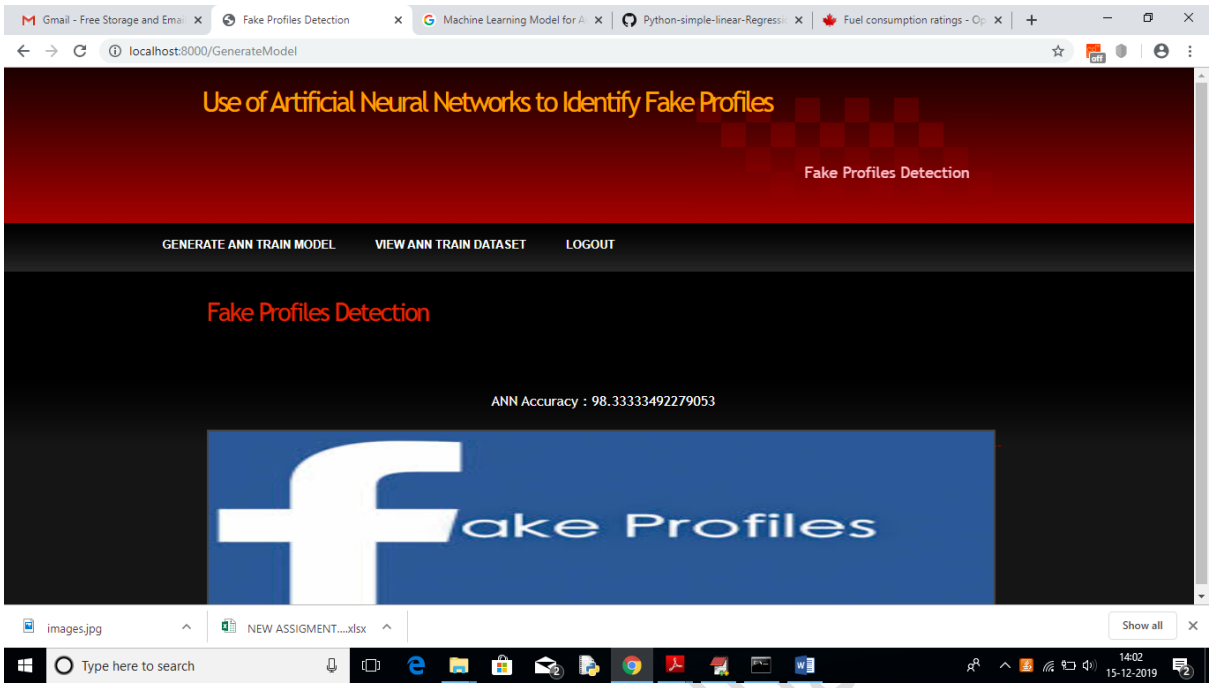
4.RESULTS AND DISCUSSIONS



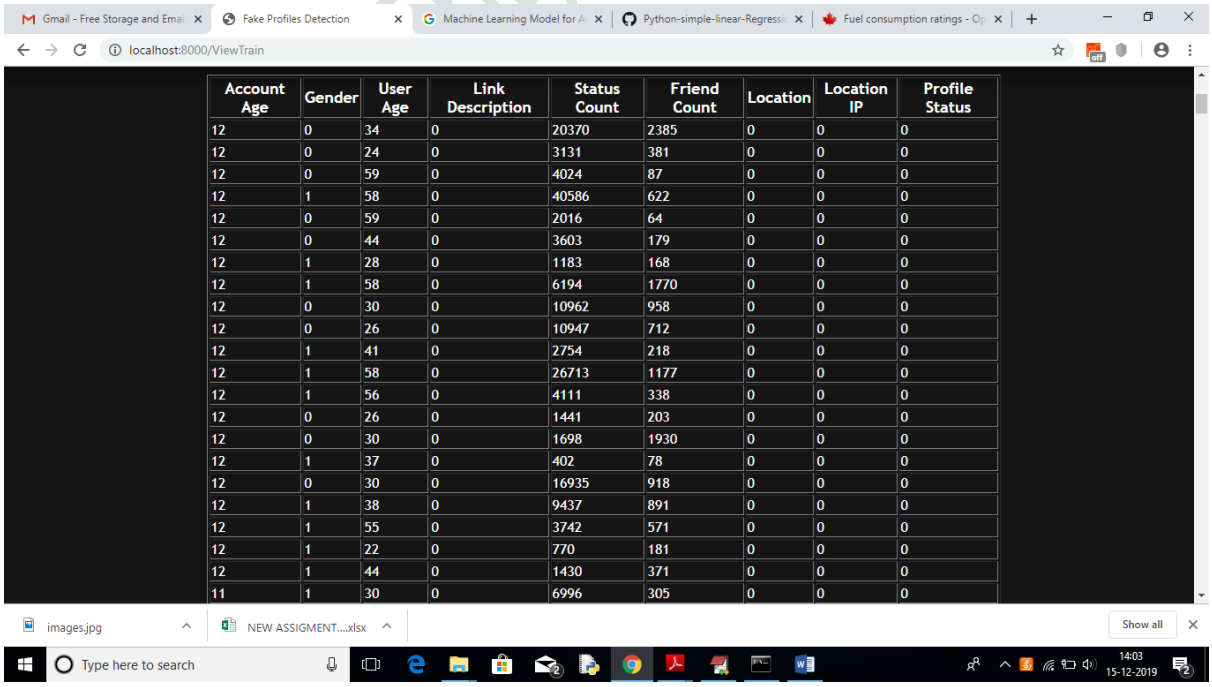
In above screen click on 'Generate ANN Train Model' to generate training model on dataset. After clicking on that link you can see server console to check ANN processing details with accuracy



In above black console we can see all ANN details.

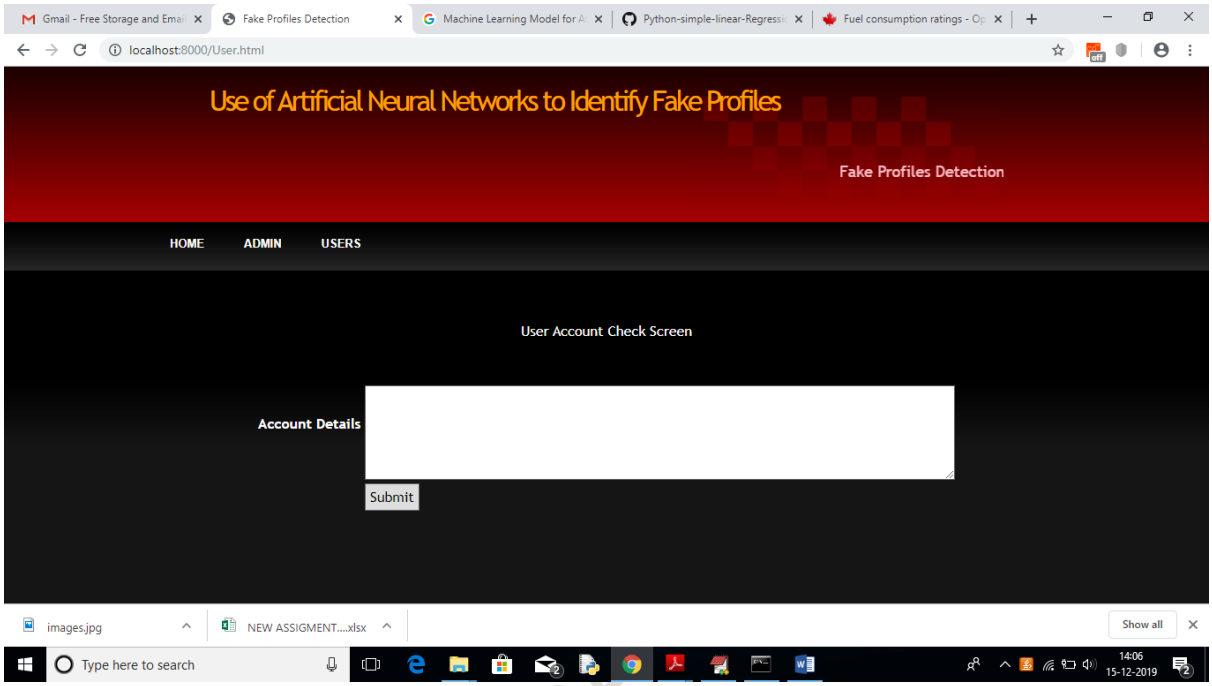


In above screen we can see ANN got 98% accuracy to train all Facebook profile. Now click on 'View Ann Train Dataset' link to view all dataset details



Account Age	Gender	User Age	Link Description	Status Count	Friend Count	Location	Location IP	Profile Status
12	0	34	0	20370	2385	0	0	0
12	0	24	0	3131	381	0	0	0
12	0	59	0	4024	87	0	0	0
12	1	58	0	40586	622	0	0	0
12	0	59	0	2016	64	0	0	0
12	0	44	0	3603	179	0	0	0
12	1	28	0	1183	168	0	0	0
12	1	58	0	6194	1770	0	0	0
12	0	30	0	10962	958	0	0	0
12	0	26	0	10947	712	0	0	0
12	1	41	0	2754	218	0	0	0
12	1	58	0	26713	1177	0	0	0
12	1	56	0	4111	338	0	0	0
12	0	26	0	1441	203	0	0	0
12	0	30	0	1698	1930	0	0	0
12	1	37	0	402	78	0	0	0
12	0	30	0	16935	918	0	0	0
12	1	38	0	9437	891	0	0	0
12	1	55	0	3742	571	0	0	0
12	1	22	0	770	181	0	0	0
12	1	44	0	1430	371	0	0	0
11	1	30	0	6996	305	0	0	0

In above screen we can see all train data and scroll down to view all records. Now ANN train model is ready and you can logout and click on 'User' link to get below screen.



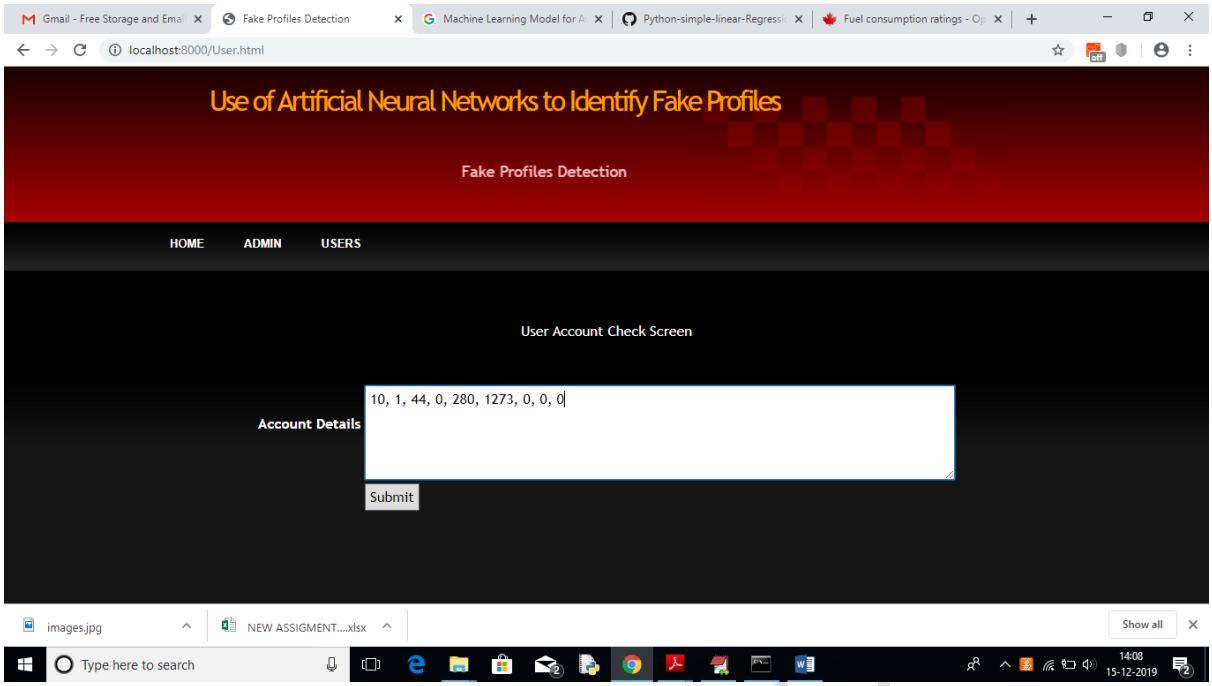
In above screen enter some test account details to get prediction/identification from ANN. You can use below records to check

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

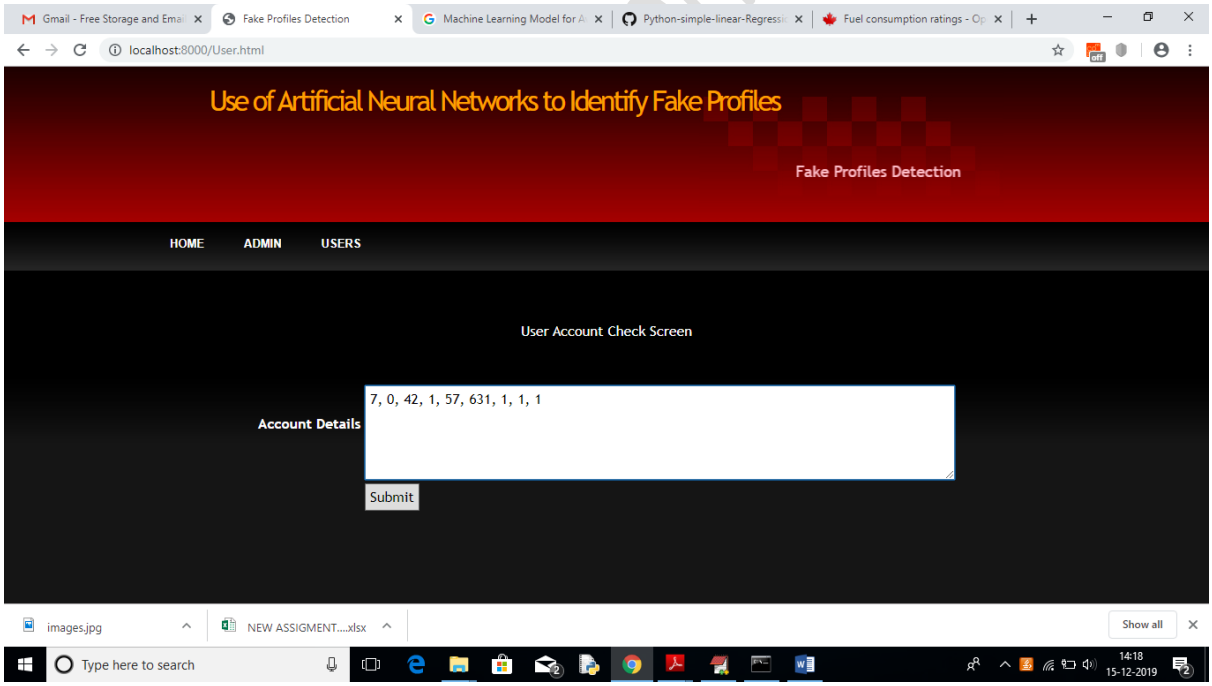
10, 0, 54, 0, 5237, 241, 0, 0

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

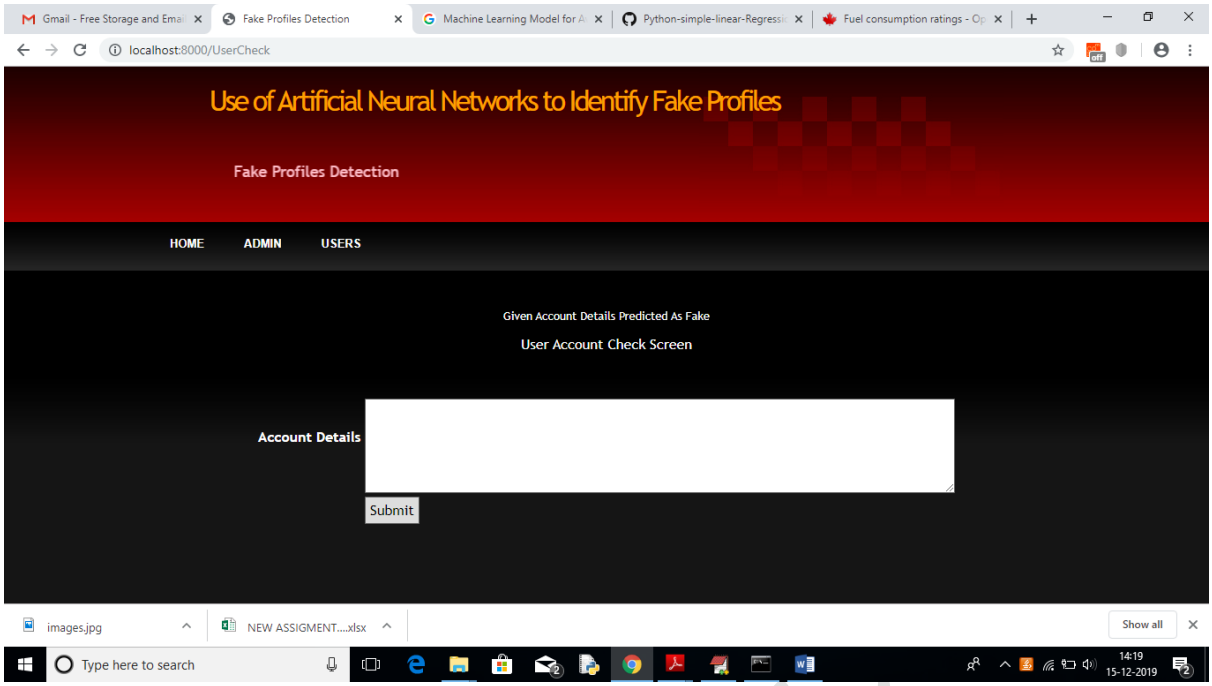
7, 1, 56, 1, 66, 623, 1, 1



For above input will get below result



For above account details we got below result



In above screen we got result as fake for given account data

5.CONCLUSION

In this paper, we use machine learning, namely an artificial neural network to determine what are the chances that a friend request is authentic are or not. Each equation at each neuron (node) is put through a Sigmoid function. We use a training data set by Facebook or other social networks. This would allow the presented deep learning algorithm to learn the patterns of bot behavior by backpropagation, minimizing the final cost function and adjusting each neuron's weight and bias. In this paper, we outline the classes and libraries involved. We also discuss the sigmoid function and how are the weights determined and used. We also consider the parameters of the social network page which are the most important to our solution

Future Work

Each input neuron would be a different, previously chosen feature of each profile converted into a numerical value (e.g., gender as a binary number, female 0 and male 1) and if needed, divided by an arbitrary number (e.g., age is always divided by 100) to minimize one feature having more influence on the result than the other. The neurons represent nodes. Each node would be responsible for exactly one decision-making process

6.REFERENCES

[1]https://www.statista.com/topics/1164/social-networks/

[2]https://www.cnbc.com/2018/01/31/facebook-earnings-q4-2017- arpu.html
[3]https://www.cnet.com/news/facebook-breach-affected-50-millionpeople
[4] https://www.facebook.com/policy.php
[5] Qiang Cao, Michael Sirivianos, Xiaowei Yang, and Tiago Pregueiro. 2012. Aiding the detection of fake accounts in large scale social online services. In Proceedings of the 9th USENIX conference on Networked Systems Design and Implementation (NSDI'12). USENIX Association, Berkeley, CA, USA, 15-15.
[6] Akshay J. Sarode and Arun Mishra. 2015. Audit and Analysis of Impostors: An experimental approach to detect fake profile in an online social network. In Proceedings of the Sixth International Conference on Computer and Communication Technology 2015 (ICCCT '15). ACM, New York, NY, USA, 1-8. DOI: https://doi.org/10.1145/2818567.2818568
[7] Devakunchari Ramalingam, Valliyammai Chinnaiah. Fake profile detection techniques in large-scale online social networks: A comprehensive review. Computers & Electrical Engineering, Volume 65, 2018, Pages 165-177, ISSN 0045-7906, https://doi.org/10.1016/j.compeleceng.2017.05.020.
[8] https://www.enigmasoftware.com/top-20-countries-the-most-cybercrime
[9] pages.cs.wisc.edu/~bolo/shipyard/neural/local.html
[10] https://stackoverflow.com/questions/40758562/can-anyone-explain-mestandardscaler [11] https://pandas.pydata.org
[12] https://www.tutorialspoint.com/python_pandas/index.htm
[13] http://www.numpy.org
[14] https://www.mathworks.com/products/matlab.html
[15] http://www.deeplearning.net/software/theano/
[16] https://scikit-learn.org/stable/
[17] https://keras.io
[18] https://www.tensorflow.org

Vol 13, Issue 02, FEB /2022

ISSN NO:0377-9254

Author's Profile

Rajasekhar Nennuri
rajasekharnennuri@gmail.com
Assistant Professor
Dept. of Computer Science Engineering
INSTITUTE OF AERONAUTICAL ENGINEERING
Dundigal, Hyderabad - 500 043,
Telangana, India.

Nisha Kumari
nk345493@gmail.com
Batchelor's in Technology
Dept. of Computer Science Engineering
INSTITUTE OF AERONAUTICAL ENGINEERING
Dundigal, Hyderabad - 500 043,
Telangana, India.

Nikitha Reddy Donti
nikhithad2000@gmail.com
Batchelor's in Technology
Dept. of Computer Science Engineering
INSTITUTE OF AERONAUTICAL ENGINEERING
Dundigal, Hyderabad - 500 043,
Telangana, India.

Naga Bharan Kaza
Nagabharan246@gmail.com
Batchelor's in Technology
Dept. of Computer Science Engineering
INSTITUTE OF AERONAUTICAL ENGINEERING
Dundigal, Hyderabad - 500 043,
Telangana, India.