CORRECTION OF MIMO CHARACTERIZATION SYSTEM LEVEL OF 5G MICRO BASE STATION USING WSN

N.Suma [1] Prof/Head Department of ECE, Dhanalakshmi Srinivasan College of Engineering Coimbatore. Email id: sumasivaravi@gmail.com
N.Bamakumari [2] Assistant Professor Department of ECE, Dhanalakshmi Srinivasan College of Engineering Coimbatore Email id: bamakumari@gmail.com
R.Siva kumar [3] Prof/Head Department of ECE, Dhanalakshmi Srinivasan College of Engineering Coimbatore Email id: Sivakumar @gmail.com

Abstract: Social Networking is the main era of data transmission as well as data creation in a large scale and the reason why big data is created is very much known. Social networking is the main platform where tons of data is being created and by 2025 even Google data centers can't handle that kind of huge volume. Increase in the fake accounts are creating exponential growth of the volume and in this paper we are proposing architecture for identifying the fake accounts in the social networking, especially in Facebook. In this research we are using Machine Learning to implement better prediction on identifying the fake account based on their posts and status on their social networking walls. We can consider Facebook and Twitter for this research and for the security purpose and availability of data, we are considering Twitter for this research. Twitter tweets and tagging, type of posts they are creating and sometimes some account people creates mess in the society with their posts, identify those and block the fake and unwanted information circulating over the network for the peace and security. Twitter is used as main based with ML NLP for the processing of the text data and here we have to use sentiment analysis for identifying the goals we put on.

*Keywords***: Machine Learning, Sentiment Analysis, NLP, Twitter, Big Data, Prediction**

I.   INTRODUCTION

 Text classification in machine learning will teach us the security levels we need to maintain in the social networking and also in our daily basis. In this social networking era we are using tweets to express our feelings and ideas or opinions in the portal through our network and as per an NGO survey, there are
 more fake accounts being created in the social networking and because of those some of the fake information is being circulating in the portals through un proper channels. In this kind of situation we need to avoid those un wanted and harmful accounts from the social networking to save space in the data centers as well and to stop the mess in the society on the social and political issues. Here in this article we are using twitter data collected from public repository as well as our own creation from a survey conducted and every day at least 1000 fake accounts are being created in the social networking and this will create burden on the network and also high volume space is being ruined. To avoid this we need to identify a path for the better implementation of the NLP and the text classification. More than NLP we can use SVM for this kind of text classification as this can be considered as the parent algorithm for the text classifications and the data identification in between the texts we wrote on the social

networking. Here we implement it with the BOW (Bag of Words) and it's a common procedure in SVM and CNB classifiers. In CNB classifiers as we use BOW and also we have to implement this on as platform like WEKA. This tool of machine learning will take the test data enven though if we are not proving the training data it will perform prediction models and give the accurate results. As per our work we have 98% accuracy of SVM to identify the fake accounts and helps to warm the social networking in charges regionally to remove and report the un wanted accounts and also CNB given accuracy of 97% which is a little less and the thing here is a criteria that bothers us is this kind of research is must be implemented in social platforms like Facebook rather than Twitter.

As per a survey there are 500 Million Female accounts in Facebook, but the population of Female in world is nearly 300 Million. That means we can assume how many fake accounts and unwanted data is being created. And how we can identify a social account is fake. It's not completely possible on identifying with status, tweets, and posts. We have to go little deeper into the concept and have to investigate into the account. We have some of the

concepts and architectures explained in this paper which are we very need to follow before creating a social network account and also how to identify the existing accounts trustworthy. [1]

In this paper we are trying to explain the existing approaches in the social networking on which they are concentrating on identify the better prediction model. In the next section we will explain the existing scenarios using in the social networking implementation, propose approach in the later section, architecture we would like to follow in the next one, results and explanation in the later section, finally conclude with conclusion and references.

II.   EXISTING SYSTEM

In the recent era of technology as the applications and the utilizations increase on our daily life, we continuously posting some unwanted and unaware stuff in the social networking and creating mess in the social platform. In the issue we can consider fake accounts and the people with fake accounts using this platform for their money and how we can stop in this kind of serious situation. Let's consider few social networking platforms for our research work. First is twitter. As we now that twitter mining is the well known application of research and we can have the public dataset of the twitter mining. In this approach we need to get the information about the people who are posting some unwanted things and they need be caught and have to remove those before it creates problem. [3-10]

The twitter account will ask some email ID or the phone number to create an account and for this kind of fake accounts people are creating fake mail accounts and we don't need to produce any authentication for creating mail Id expect phone number. If you are following any improper things in the social networking we need to identify that before spreading to others. Some issues like country problems. A political drama can be added as a tag in the twitter with and Hash tags. If the hash tag is not created by the person or account which is not responsible then that will create a social war in the network and they used completely discussing about this with the bad words and the comments. Here we are focusing on the comments too for identification of the information and the sample hash tagging

scenario will be explained with the image below. In the image below that is Figure 1 illustrates that the hash tagging in the twitter will go on continuously with unwanted stuff. Once tag is created with the another tag is created by another person and for a limit of tweets or the likes for the concept in the tag twitter will pay some amount for the account holder. Including money fake information may spread through it and this causes a big problem.[2-7]

The main problem here is we are not considering the text in with which tag is applied. First we need to identify the whether the tag applied is genuine information of not. If the tag is genuine information then we can make that tag popular. If not we need to spam or report that tag to avoid mess in the social platform. How to identify that a tag is genuine or not. Based on the previous tags and comments of that person who is holding that account. If we track all the tweets and comments of the person who posted that hash tag we can predict that whether that tag can be published or not. For this kind of things we need to discuss the procedure and we will do the same in the proposed approach section which will be explained in the next few lines.[5-13]

Another social networking we consider its Instagram. Instagram is very famous for following celebrities and the fan pages in our own locality or globally. Day by day fake accounts are being increased in instagram more than facebook. As this was invented by Facebook, more than facebook Instagram is more popular for creating fake news and creating reportable content. If we observer Instagram for a while, we have some many women who are doing unfair things for money through this platform. They used post some unfair words in their instagram story and that story will be appear for 24hours from the time of publish. The current architecture of Instagram is not bothering about the posts and story of the account holders until anyone reports a live video call or a post. We need to identify the words while the account holder posting and we need to identify the sentence meaning in that post. As a dignity purpose we are not mentioning those sentences and words in this article and if they are required further more for any explanation and proof purpose anyone contact use we are very happy to explain which kind of words we need to consider in instagram for

indentifying fake accounts. Figure 2 will explain the architecture of Instagram posting and Figure 3

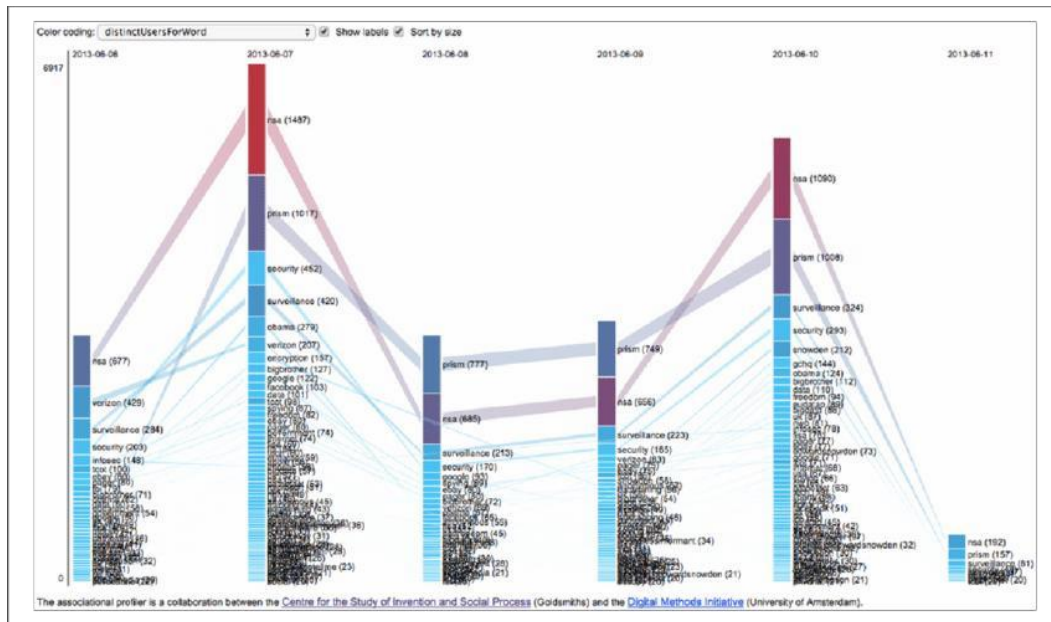explains the database architecture models for the instagram[10-15].
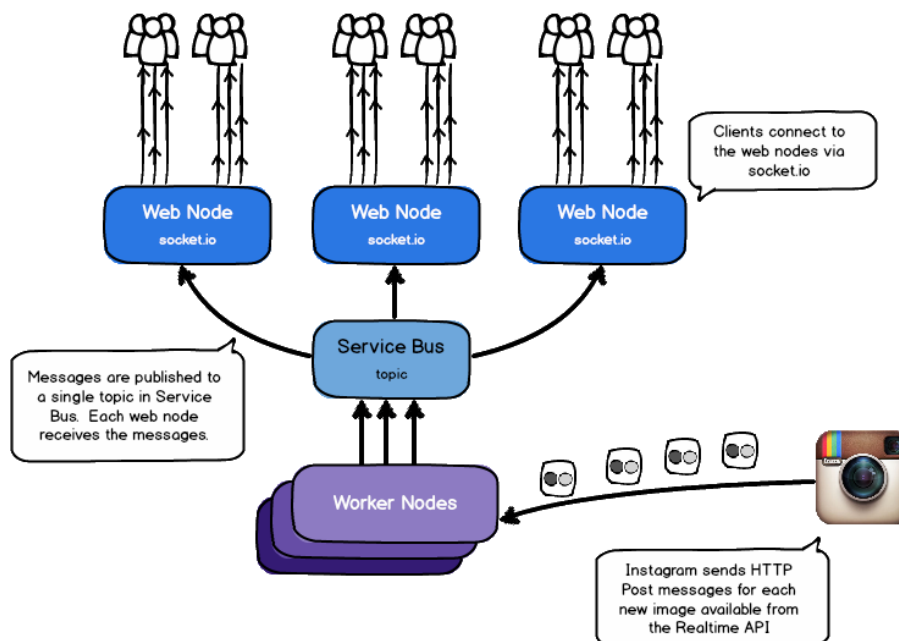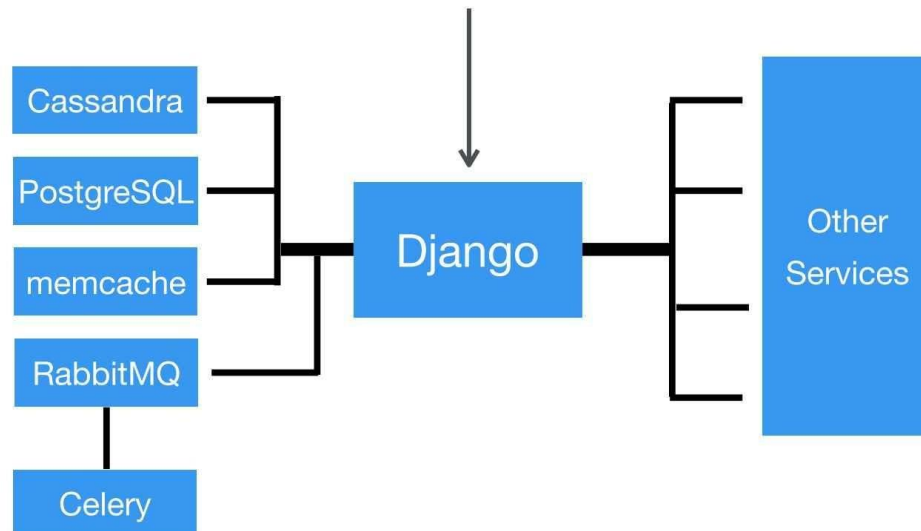
Figure 1: Twitter Hash Tagging Scenario

Figure 2: Architecture of Instagram following for posting with nodes

Figure 3: Instagram Stake Architecture for Database Maintenance.

In this architecture we need to observe that how data is moving from each edge to other edge. In this architecture when the post is submitted by the person it will navigate to the worker node in the server. In the real time for each image or the text in the instagram story instagram app will send the HTTP request to each image which are available from real time API. These API's are created or when can use existing image handling API's. In this kind of architectures we have multiple working nodes available and if you don't have any working node unavailable then the posting will be delayed or aborted with time out. The point here is we need to concentrate is when the data is posting through HTTP request to the worker node, we are not validating the content in the post. Whether that post may cause harm to the society, or some nudity is available, something related to social issues like prostitution or something related to these. In this case we are failed in development of client secured applications though we have high range of servers to maintain and manage.[14-18]. The very second example we need to put in front of the technocrats is unhealthy twitter tags. Improving political and general awareness is a good thing but this kind of unhealthy hash tags will cause damage to the human behavior this causing some of the quarrels[16] among the communities. Instead of using social network to stop the issue, some fake account holders will create issue with their comments and tags. Our machine learning approaches currently using failed to identify those before publishing the post.[18] Third issue is Fake accounts in Facebook. As I mentioned earlier, there are more than 200 million Fake Facebook accounts in the name of female. That's a great thing that Facebook is not controlling their accounts and because of these fake accounts money laundering, smuggling and some other anti-social elements are taking place[17]

In the next section we will address the proposed architecture for this validations before posting an status and secondary thing is identifying gender based on the content in the post.

PROPOSED APPROACH

In this approach we are focusing on an architectural methodology for implementation like below. Which will have few stages in their implementation and the concept here is to predict the Fake accounts in the social networking and also try to avoid the fake status or contents in the social media[18].

a. *Data Identification*

Data identification is a great task for social media predictions as this is one of the sensitive areas to discuss. Because of some security issues we are not providing the information abut the public repository where we got the dataset with the relevant information. But we are here with the data needed to

identify whether the account is fake or not and the secondary thing is to identify the fake news and block that information to circulate in you locality.

The data we need here is the account holder general information, age given to create account, what is the most used word in their post, are there any contents related to nudity, are there any content of abuse on a community or a gender etc.

In this paper we are proposing two kinds of individual architectures and finally we will collaborate two concepts and to make a final product. First one will identify the account details by using NLP and network identification. Based on the network on which two or more accounts are accessing then we need to ask them to provide security verifications. That will be explained in this architecture easily.
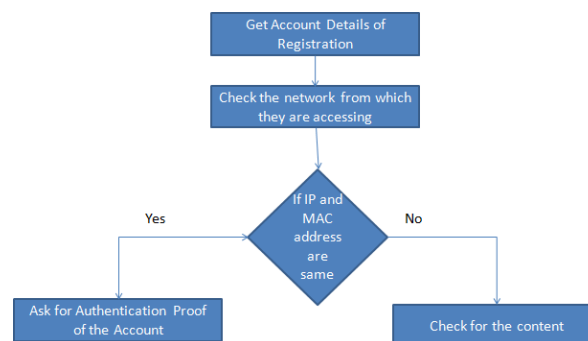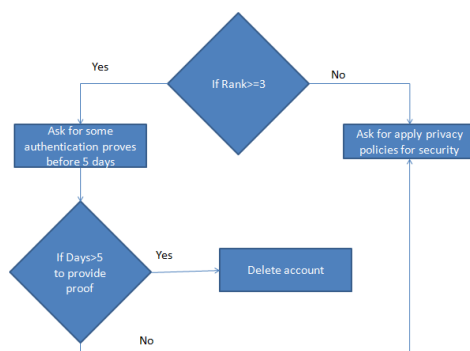

Figure 4: Architecture 1 for account authentication
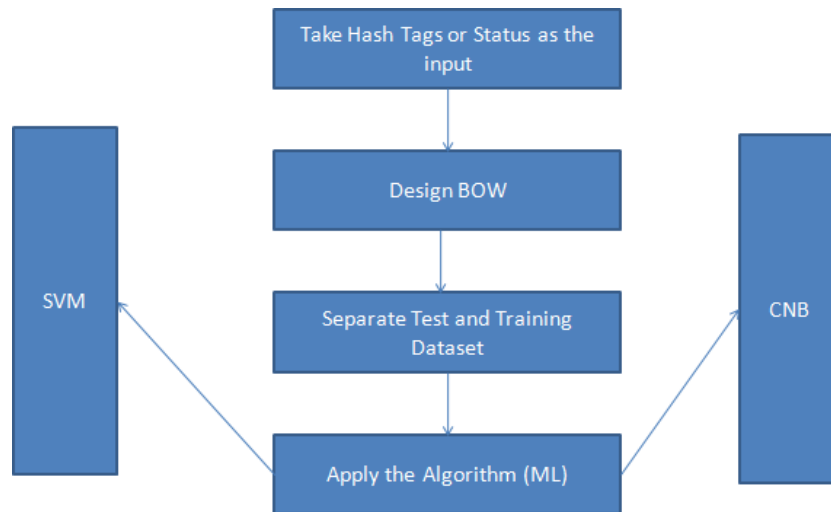
Figure 5: Architecture 2 for Authentication before using ML algorithms in complete way
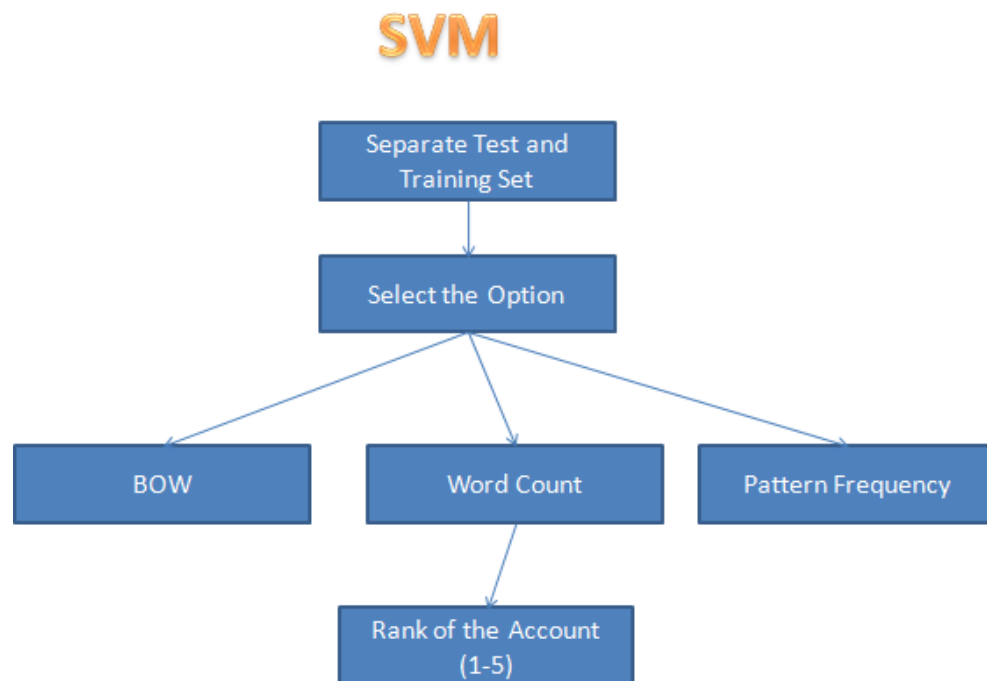
**Figure 7: SVM architecture 2 for Authentication**

In the above scenario of architectures we have divided into four parts, which is two parts each in each type. In first architecture part one in figure 4 the client using the account have to face some of the security issues like if two or more accounts are managing from one IP address or MAC address, then our application will ask for authenticate themselves with ID proof. This is because we don't know the same person with different accounts is managing from the same address. If the same is like this " same person is using two or more accounts and if the one is from desktop and other two is from mobile device. For these kind of situations based on the common IP address and the name relativity of the account need to authenticate and if there is any security bleach we need to ask for provide proof for further continuation with the account. If the IP address is different and we can check with MAC address from which the account is getting access using the server log files and request and response times.In the second part of the first architecture we are ranking the account with this kind of security bleach and if the rank is 1 or 2 there will be no problem and ask them for apply security privacy settings to protect their account. If the rank is greater than or equal to 3 then we need to ask for the proof.

If the person with multiple accounts failed to provide proofs within 5days then account and the data of that account must be terminated for the security purpose without any intimation. If they provided proofs, then ask them to apply privacy settings after log in into the account. In the second architecture we are using SVM BOW concept for identification of the number of words are harmful. The harmful words are collected and gathered as the data set or in the text format. We need to divide the test and training dataset and attach the BOW to the prediction model through SVM. It will calculate the number harmful words are there in individual account and this is based on their content identifying account user and warning him or her to provide the authentication proofs to continue with the account further. In this process we will identify the word count of the single word, number of times a pair of words that are harmful are repeated. So that if the scale is less than 3 then the user is fine, if the scale is greater than or equal to 3 then the user must provide proofs to continue with the account. This process can be useful for any social networking account.The below things will explain the data cleaning we need to do before processing any data set .
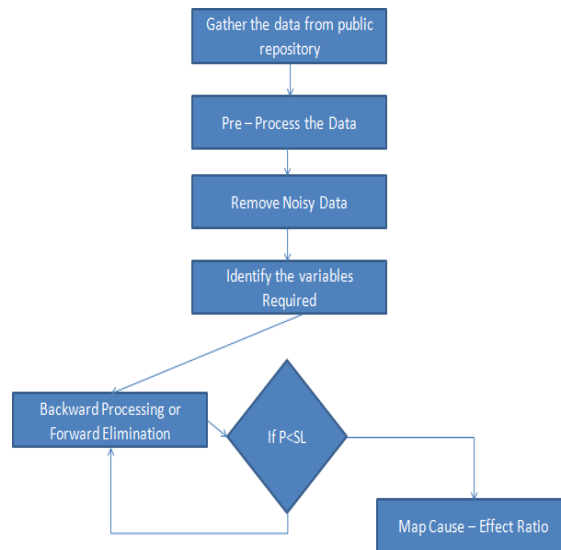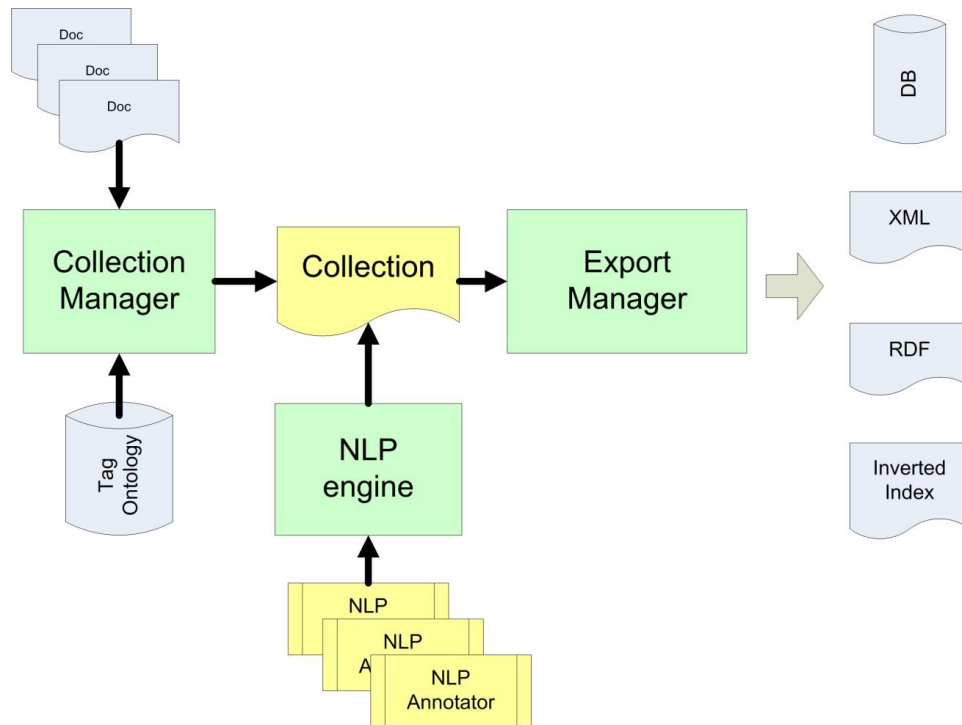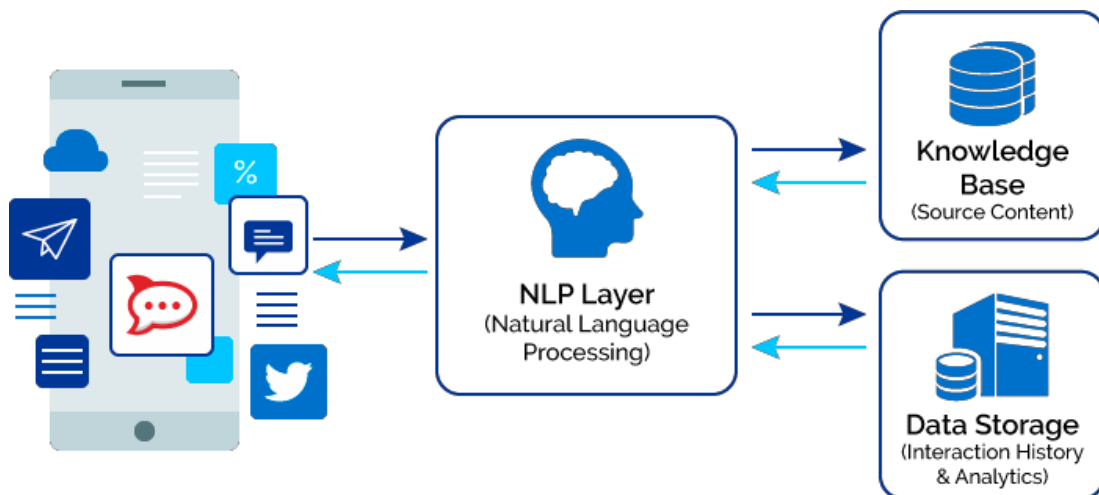
Figure 8: Part 1 of the Propose Architecture
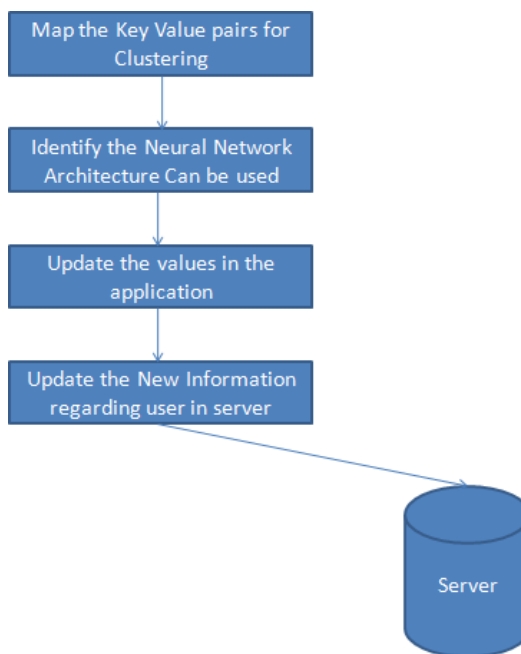
Figure 9: Part 2 of the proposed architecture

Figure 10: NLP Architecture to use for concerting some of the unknown language types to understandable format

Figure 11: NLP architecture which will maintain the data from application to repository

RESULTS

In this section we will discuss about the clusters we formed and in the table we have key value pairs which will have the group of pairs. This will tell the medication limit of the patient based on the age, gender, weight, previous treatments, previous medication, and current scenario. The below is the table which will give the explanation based in the key value pair. This pair will give the warning to the doctor if her give the high drug usage.These are the few clustering results we acquired as mentioned below which will have the values of how many times a pair of key value is being repeated.
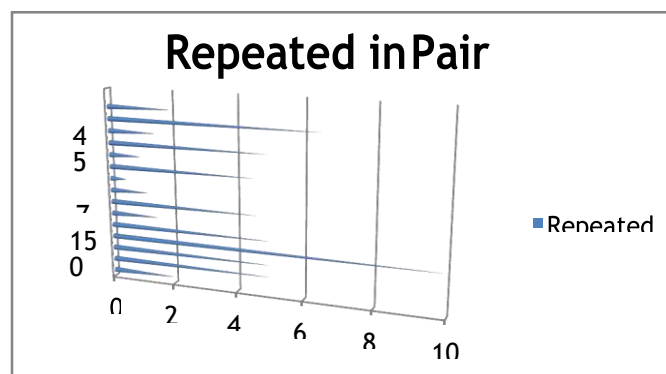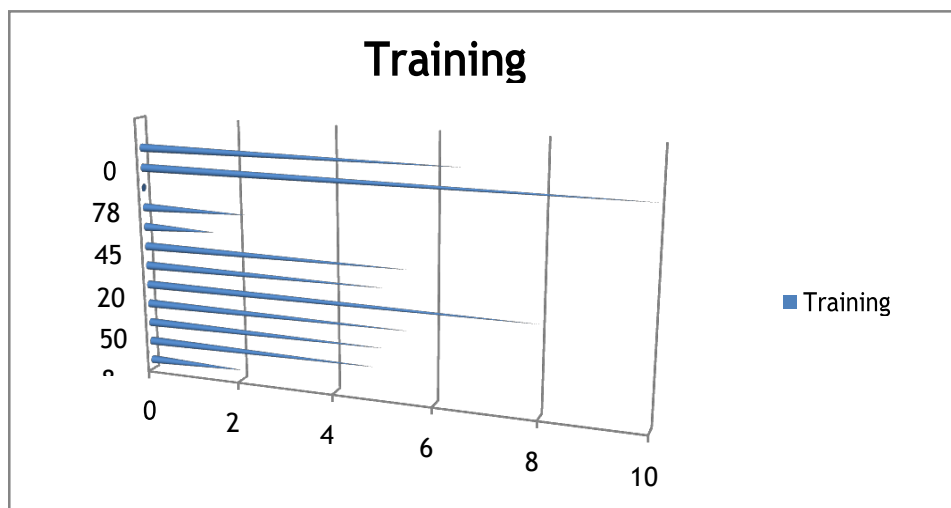

Figure 12: Key Value pair sets occurrence result

Social Networking type	SVM Accuracy	CNB Accuracy	Number of accounts processed	No . Of Valid Accounts	No. Of Invalid accounts	No. Of accounts deleted
FB	97	95	15000	12000	3000	1500
Instagram	95	95	15000	12000	3000	1000
Twitter	99	97	15000	12000	3000	1000
YouTube	99	98	15000	12000	3000	500
Whatsapp	89	89	15000	12000	3000	1500

CONCLUSION

The main issue with social networking security is not authenticating them properly before publishing the data. Here we used some of the chats, status and all the account information and proposed an architecture suing which we need to identify the genuineness of the account so that based on which we can make that account continue with the service or we need to terminate the service. SVM and CNB are used in this process for validating the content based on the text classification and sentiment analysis of the data. In this sentiment analysis we need to gather the harmful words count, how many times they are repeating, and what is the harmful pair of words and how many time they are repeating. For Facebook SVM shows 97% accuracy as CNB shows 95% percent accuracy of identifying the fake accounts.

References

1. A. Graves, et. al. A Novel Connectionist System for Improved unconstrained Handwriting Recognition.IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 31, no. 5, 2009.
2. Ciresan,Dan, Ueli Meier,and Jurgen Schmidhuber. "Multi-column Deep Learning Neural Networks forImage Classification.IEEE Conference on Computer Vision and Pattern Recognition. (2012): 3642-649.IEEE. Web.
3. Hsu, Chih-Wei, and Chih-Jen Lin. "A Comparison of Methods for Multiclass Support Vector Machines.".IEEE Transactions on Neural Networks. 13.2 (2002): 415-25. Web.
4. Knerr, S., L. Personnaz, and G. Dreyfus. "Single- layer Learning Revisited: A Stepwise Procedure for Building and Training a Neural Network."Springer. 68 (1990): 41-50. Web.
5. Liu, Cheng-Lin, Kazuki Nakashima, Hiroshi Sako, and Hiromichi Fujisawa. "Handwritten Digit Recogni-tion: Investigation of Normalization and Feature Extraction Techniques.". Pattern Recognition. 37.2 (2004):265-79. Web.
6. M. Chen, Z. Xu, K. Weinberger, and F. Sha. Marginalized denoising autoencoders for domain adaptation. In Proceedings of the 29th International Conference on Machine

Learning, pages 767{774. ACM, 2012.

7. G. E. Dahl, M. Ranzato, A. Mohamed, and G. E. Hinton. Phone recognition with the mean-covariance restricted Boltzmann machine. In Advances in Neural Information Processing Systems 23, pages 469{477,2010.

8. O. Dekel, O. Shamir, and L. Xiao. Learning to classify with missing and corrupted features.Machine Learning, 81(2):149{178, 2010.

9. Globerson and S. Roweis. Nightmare at test time: robust learning by feature deletion. In Proceedings of the 23rd International Conference on Machine Learning, pages 353{360.ACM, 2006. [10]I. J. Goodfellow, D.Warde-Farley, M. Mirza, A. Courville, and Y. Bengio. Maxout networks. In Proceedings of the 30th International Conference on Machine Learning, pages 1319{1327. ACM, 2013. [11]G. Hinton and R. Salakhutdinov. Reducing the dimensionality of data with neural networks.Science, 313(5786):504 { 507, 2006.

10. G. E. Hinton, S. Osindero, and Y. Teh. A fast learning algorithm for deep belief nets.Neural Computation, 18:1527{1554, 2006.

11. M.Arun Kumar and N Kirthika, ”Efficient implementation of ROM-less FFT/IFFT processor using fused multiply and added unit,” International Journal of Electrical and Electronic Engineering & Telecommunications, vol. 2,No. 2,April 2013.

12. K. Jarrett, K. Kavukcuoglu, M. Ranzato, and Y. LeCun. What is the best multi-stage architecture for object recognition In Proceedings of the International Conference on Computer