Knowledge Based Sentimental Analysis and Deep Learning

Sheena Popli[1], Dinesh Harkut[2]

[1]*Computer Science And Engineering, SGBAU,*
[2]*Computer Science And Engineering, SGBAU,*

Abstract— Popularity of Online Social Networking sites is increasing day by day. It allows users to connect with each other and share their views and opinions. Today, Online Social Networking (OSN) has become a universal platform where users can openly share their feelings and expressions with each other. Usually, users express their sentiments and emotions through the messages they post on OSN. Some messages indicate that the person is happy and some indicates that the person is sad according to the words present in the sentence. It has become important to detect users with psychological disturbances such as stress and depression and send them calm, motivating messages in order to improve their emotional state. This can be done with the help of sentiment analysis and deep learning.

Keywords— Sentiment Analysis, deep learning, online social networking, recommendation system, recurrent neural network .

I.   INTRODUCTION

Online social networking sites have exploded in popularity. The number of active online social network (OSN) users has grown considerably. This high number of users, on OSN, is mainly due to the increase of the number of mobile devices, such as smartphones and tablets, connected to the Internet. Currently, OSN have become a universal place where opinions, feelings, views can be openly shared. In recent years, the analysis of the messages posted on OSN has been used by many applications such as industry of health care informatics. The sentiments that are expressed on the messages posted by the user on OSN provide various clues with respect to the behaviour of users. For example, sentences containing positive words may show that the person is happy and calm. On the other hand, sentences containing negative words may show that the person is sad or stressed. If the level of sentiment intensity value remains low or it changes from high to low, then this can indicate that the person is emotionally disturbed. Depression is one of the most common disorders found all over the world. Unfortunately, it is not always easily recognized. Sentiment Analysis includes the use of Natural Language Processing to identify, extract and study different states of information. One of the basic tasks of sentiment analysis is to identify the polarity of the text- whether the text in a document or sentence is positive, negative or neutral.

II.   LITERATURE REVIEW

In survey, numerous researchers contributed and discussed about various techniques of sentiment analysis, few important in the context of our problem are shared in this review paper.

[1] *M. Al-Qurishi et al.* represented a study on proposed approach to create an integrated social media contents analysis platform that leverages three levels of features to analyze and find out the anomalous behavior that deviates significantly from the norm in large scale social media. For this, various number of user profile are gathered from twitter and YouTube along with almost 13 million channel activities. Various evaluations also get conducted.

[2] *H. Lin et al.* finds that user's stress state is totally related with the friends on social media. Here, a large dataset is gathered from real world social platforms to study the stress state and social interactions of user. A novel hybrid model is created which contains stress related attributes like textual stress, visual stress and social stress. For analyzing the stress the concept of sparse connection and delta connection is used.

[3] *G. Sannino et al.* proposed an approach that requires a photoplethysmography sensor and a mobile phone. It uses genetic programming to automatically find a relationship between blood pressure value and photoplethysmography value. This is tested based on eleven subjects and compared with other regression method.

[4] *A. E. U. Berbano et al.* presented further work on neural engineering that aims on the classification of mental, emotional, physical and no stress with the help of EEG signal analysis. It is very important to detect people with stress. Stress is detected through EEG signal analysis. The results are then pre-processed, extracted and selected with the help of Discrete Wavelet Transform (DWT). This is then used as an input to detect stress using Artificial Neural Network and tested using K-fold cross validation method.

[5] *J. Ham et al.* states that nowadays people are suffering from various stress conditions. The stress can cause various symptoms and in some cases it can lead to dangerous disease. That is why, it is necessary to develop to a method that evaluates stress. HRV (Heart Rate Variability) of normal and healthy people is analyzed by using

photoplethysmogram (PPG) while using various stress situations. The extracted HRV features and Linear Discriminant Analysis (LDA) are used in order to detect and classify the stress levels.

[6] *S.Tsugawa et al.* represented the effectiveness of using social media activities of a user which is evaluated for finding the degree of depression. Here, several features are extracted from the activity histories of users. Using these features, the model is constructed to estimate the presence of depression in activities. These features can be used to detect depression from the activities of users.

[7] *R. G. Guimar˜aes et al.* defines a root mean square method in which age group is considered. The characteristics of different age groups, classification of different age group- such as teenage, adult age are considered. On the basis of these groups, different machine learning algorithms are tested and classified. Sentiment analysis is performed using this classification. For this, a detailed analysis with 7000 sentences is performed to determine whether the characteristics are relevant or not.

[8] *G. Lample et al.* defined the concept of deep learning which is used to obtain highest performance without using any specific feature or gazette. NER (Name Entity Recognition) is very challenging problem for languages with low resources. Here the concept of Bi- directional LSTM is used. In this, word and character level features are automatically detected and the need for most hand- engineering features is eliminated. This model obtains highest score in languages like Persian. It can also be applied for datasets.

[9] *R. Rodrigues et al.* proposed a sentiment analysis tool named SentiHealthCancer. This tool improves the detection of emotional state of patient. This tool is used in Brazilian Online Cancer Communities by checking their posts written in Portuguese language. This tool is basically used to detect positive, negative and neutral messages of patients in online communities of cancer patients.

[10] *S. Poria et al.* presented a method of identifying the sentiment polarity of people speaking in videos. They collected different evidences such as the facial expressions, the words they said as well as the sound in the speech. Deep CNN is used to extract features from text and MKL is used to classify the heterogeneous fused feature vectors. Combined feature vectors of textual, visual, and audio modalities is used to train a classifier which is based on multiple kernel learning.

[11] *Yoon Kim* described a series of various experiments with CNN built on word2vec. They showed that a CNN with static vectors and little hyper parameter tuning could achieve great results. A simple modification was made that allows to use both static vectors and task-specific vectors.

[12] *Y. Chen et al.* proposed a recommendation system that recommends an effective emotional state for the person with negative mental state with positive state. Contrast pattern mining is used to find the common contrast behaving way of happy and sad users. Behaviours and habits of the users are used to find these patterns. A dataset of people with diabetes or heart disease is used in this work. This word is beneficial for health care area.

[13] *J. Tang et al.* proposed a quantitative research on how emotional state of a user can be found by using the history log of emotions and how the user's psychological state is influenced by its friends on social media. They used a method called MoodCast for determining emotion dynamics in social network. In this work Metropolis-Hastings algorithm is used to obtain a solution.

[14] *Y. Xue et al.* stated that adolescents face a lot of psychological pressure due to study, work, affection etc. If these problems are not cured, it could lead to mental illness. Micro- blogging has become a popular media channel for teenagers. They investigated a lot of features from the their tweets and tested them on 5 classes such as Naïve Bayes, Random forest, Support Vector machines, Artificial Neural Networks and Gaussian Process Classifier for pressure detection. They showed that Gaussian process Classifier has the highest accuracy in detection due to its robustness.

III. PROPOSED WORK

The main objectives of the study are listed below:

1. To develop an online Emotional health monitoring system using social networking

2. To implement sentimental analysis to detect user's mental situation

3. To implement genetic algorithm to optimize RNN

4. To implement BLSTM-RNN with genetic algorithm for the disorder entity recognition
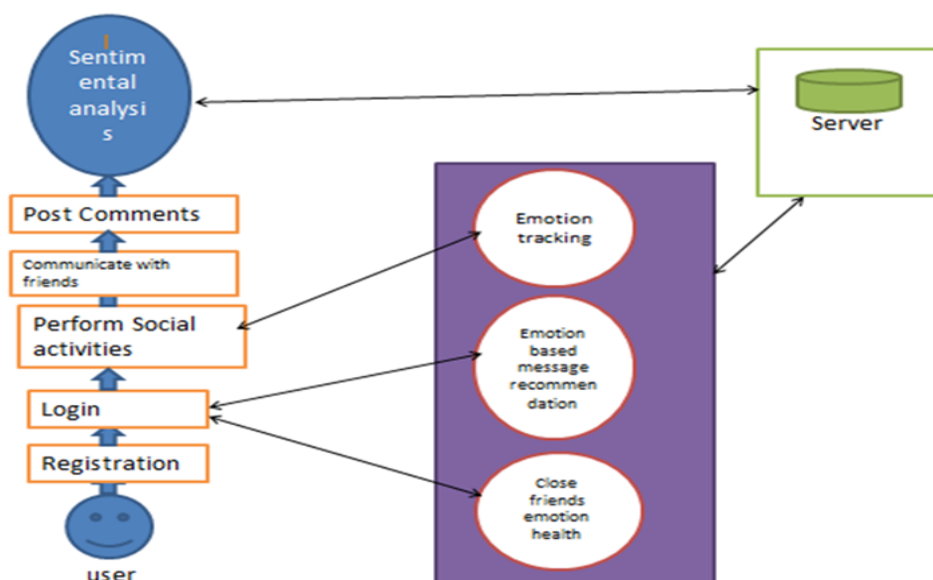
Fig. 1 proposed work Diagram

The proposed work, shown in above figure, is planned to be carried out in the following manner:

1. First the user will register itself on the social networking site by specifying his name, address, gender, mobile number, email address, working status etc. It is mandatory to select a username and password. For the security reasons, a security question and an answer for that question is required.
2. By using the username and password, the user will login.
3. The user can perform social activities such as sharing of documents and pictures with their friends. The friends will have the right to view, download or delete the shared information.
4. Communication which is the main feature of the social networking sites can also be done with the friends in the form of personal chats, posting on timeline. Sentiment analysis will be performed on the messages that are posted by the users in order to detect the emotional state of the users.
5. One extra feature will also be added in this work which will be self- communication. The user can communicate with itself by maintaining a personal diary. Sentiment analysis will be performed on this daily.
6. The polarity of the sentences posted by the users will be determined by performing sentiment analysis and the result will appear positive, negative or neutral depending on the words present in the sentences. This will be done directly with respect to the database server.
7. Based on the messages shared between the friends, close friends will be detected. Warning messages about the depressed user will be sent to close friends.
8. There will be an expert which will be registered by the admin. The expert will have the right to view the depressed entities and send them motivational, calm, happy and relaxing messages.
9. In this manner the emotional state of the depressed user will be improved.

IV. CONCLUSIONS

Various techniques of sentiment analysis are presented. After going through the various research papers, it is observed that Recurrent Neural Network (RNN) is used. These networks remember only the important things about the received input which makes them precise in predicting what's coming next. Instability is the main problem with recurrent networks. Here, the outputs are the inputs. This leads to self-reinforcing feedback cycles which gets in the way of the design goal of the network. Because there is no clear difference between dependent variables and independent variables, they are difficult to analyze. As RNN is very time-consuming, we will use Genetic algorithm to optimize it. By using genetic algorithm a lot of time will be saved.

REFERENCES

[1] M. Al-Qurishi, M. S. Hossain, M. Alrubaian, S. M. M. Rahman, and A. Alamri, "Leveraging analysis of user behavior to identify malicious activities in large-scale social networks," IEEE Transactions on Industrial Informatics, vol. 14, no. 2, pp. 799–813, Feb 2018.
[2] H. Lin, J. Jia, J. Qiu, Y. Zhang, G. Shen, L. Xie, J. Tang, L. Feng, and T. S. Chua, "Detecting stress based on social interactions in social networks," IEEE Transactions on Knowledge and Data Engineering, vol. 29, no. 9, pp. 1820–1833, Sept 2017.

[3] G. Sannino, I. D. Falco, and G. D. Pietro, "A continuous non-invasive arterial pressure (cnap) approach for health 4.0 systems," IEEE Transactions on Industrial Informatics, pp. 1–1, 2018.

[4] A. E. U. Berbano, H. N. V. Pengson, C. G. V. Razon, K. C. G. Tungcul, and S. V. Prado, "Classification of stress into emotional, mental, physical and no stress using electroencephalogram signal analysis," in 2017 IEEE International Conference on Signal and Image Processing Applications (ICSIPA), Sept 2017, pp. 11–14.

[5] J. Ham, D. Cho, J. Oh, and B. Lee, "Discrimination of multiple stress levels in virtual reality environments using heart rate variability," in 2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), July 2017, pp. 3989–3992.

[6] S.Tsugawa, Y.Kikuchi, F.Kishino, K.Nakajima, Y.Itoh, and H.Ohsaki, "Recognizing depression from twitter activity," in Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems, 2015, pp. 3187–3196.

[7] R. G. Guimar˜aes, R. L. Rosa, D. D. Gaetano, D. Z. Rodr´ıguez, and G. Bressan, "Age groups classification in social network using deep learning," IEEE Access, vol. 5, pp. 10805–10816, 2017.

[8] G. Lample, M. Ballesteros, S. Subramanian, K. Kawakami, and C. Dyer, "Neural architectures for named entity recognition," in Proceedings of the NAACL, San Diego, EUA, May 2016.

[9] R. Rodrigues, R. das Dores, C. Camilo-Junior, and C. Rosa, "Sentihealth-cancer: A sentiment analysis tool to help detecting mood of patients in online social networks," International Journal of Medical Informatics, vol. 1, no. 85, pp. 80–95, 2016.

[10] S. Poria, E. Cambria, and A. F. Gelbukh, "Deep convolutional neural network textual features and multiple kernel learning for utterance-level multimodal sentiment analysis," in EMNLP, 2015.

[11] Kim, Y, Convolutional neural networks for sentence classification. arXiv preprint arXiv:I408.5882, 2014.

[12] Y. Chen, M. L.-J. Yann, H. Davoudi, J. Choi, A. An, and Z. Mei, "Contrast pattern based collaborative behavior recommendation for life improvement," Jun 2017, pp. 106–118

[13] J. Tang, Y. Zhang, J. Sun, J. Rao, W. Yu, Y. Chen, and A. C. M. Fong, "Quantitative study of individual emotional states in social networks," IEEE Transactions on Affective Computing, vol. 3, no. 2, pp. 132–144, April 2012.

[14] Y. Xue, Q. Li, L. Jin, L. Feng, D. A. Clifton, and G. D. Clifford, "Detecting adolescent psychological pressures from micro-blog," in Health Information Science. Springer International Publishing, 2014, pp. 83–94