IJCSE International Journal of Computer Sciences and Engineering Open Access
Research Paper Volume-6, Special Issue-3, April 2018 E-ISSN: 2347-2693

Prediction of Social Media User's Mood using Deep Learning

M.K.Sandhya[1*], V.Soundarya[2], R.Swarnalakshmi[3], I.Swathi[4]

[1,2,3,4]Dept. of Computer Science and Engineering, Meenakshi Sundararajan Engineering College, Chennai, India

[*]*Corresponding Author: mksans@gmail.com*

Available online at: www.ijcseonline.org

***Abstract*—** In recent times, there is a huge increase in the usage of social media to share one's opinion, feelings and even daily activities. By predicting the mood of the users in social media, we can identify the users who discuss or express suicide-related information. Prediction of user's mood based on the likes, shares and status posted by them on social media is a challenging task as the mood of users change frequently. In this paper, a scheme is proposed to predict the user's mood based on the likes, shares and status posted in social media and identify the users in the state of depression. This scheme classifies the mood of user as happy, sad, neutral, angry etc. using deep learning. It presents news feeds to keep the depressed user happy and enthusiastic. When the user is in a prolonged state of depression, the alert system alerts the top five users in his/her friend list. This scheme predicts the mood of the users with accuracy around 87%. Further, time critical information is sent to some users who regularly share information such that it reaches all the users within a certain period of time.

***Keywords*—**Mood Prediction, Social Media, Alert System, Time Critical Information, Depression, News Feed

I. INTRODUCTION

Emotions are the way of expressing human's feelings. Nowadays, most of the people express and share their feelings via social media. Social media brings up the social interaction among people and acts as an interface for humans to express their feelings according to their mood swings. The expression of their emotions in social media has its pros and cons. The mood of the user can be classified into different types such as happy, sad, disgust, angry, contempt, etc. This mood of the user swings due to current situation, stress, clinical activities etc. The mood of the social user is predicted based on their emotions they express as likes, shares and status posted. The need to predict the mood of the social media users is to promote marketing for business development, prevent suicidal activities etc. The major advantage of using social media in business is to track the user's likes and interests and then the marketing takes place in response to user's interest. The second advantage is that it is the easiest way to communicate and get connected with all the people [1]. The drawback of the social media is that some users wrongly use this media and land up in prolonged depression leading to suicidal activities. This occurs due to the mood swings of the user. To overcome this drawback, the mood of the user is predicted using their postings. The mood swings are monitored for determining the user's characteristics. Machine learning techniques are used to classify the data based on the training sets. Machine learning is nothing but feeding the dataset to a machine and making it learn. This helps in predicting user's mood using deep learning. The statistics of Facebook [2, 3] such as the total number of users, the number of posts, likes etc. is presented in Table 1.

Table 1. Statistics of Facebook

The large number of posts in social media is extracted as datasets and it is used as input for mood prediction using machine learning techniques. These datasets can be classified as text or images based on their type. The attributes of these datasets are extracted and classified as positive and negative attributes using classifiers. These attributes play an important role in mood classification. The electronic wearable sensor devices can also be used to predict the mood of the user but these devices are either costly or may be subjected to damage. In this paper, the mood of the social media user is predicted by monitoring the activities of the user by analyzing the posts liked, shared and status posted by the user using deep learning. This scheme identifies the users in prolonged depression and alerts the top five users in their friend list.

The rest of the paper is organized as follows: Section II presents the work related to mood prediction of users in social media or in the internet. Section III the presents the proposed method of mood prediction of users in social media using Deep Learning. Section IV discusses the results of the proposed methodology. Section V concludes the research work with future directions.

II. RELATED WORK

In this section, a detailed review on the work related to mood prediction of users in social media or in the internet is presented.

Chiranjeevi et al. [4] proposed a neutral face classification technique using personalized appearance models for fast and robust emotion detection. In this work, the major challenge that is faced is robust facial recognition. Using supervised learning, the facial expression recognition is done by extracting the features or patterns or key emotion points. It cannot accommodate all expressions or emotions such as pose, lighting, facial bias, etc. It classifies emotions of the user using the key emotion points. By detecting neutral state of the user it will learn neutral appearance at key emotion points. The accuracy is maintained and the complexity is reduced. The textual statistical model like head motions and distortions are detected using key emotion points. The major drawback is that it is not implemented in social media and does not handle depression recovery.

Number of unique users monthly in Facebook	2.07 Billion
Number of Facebook posts	2 Trillion
Average number of time spent on Facebook per user per day	20+ minutes
Number of Facebook friends for average user	160
Total number of Facebook likes	1.13 Trillion
Average number of images uploaded per day	300 Million
Average number of mobile active users	1.15 Billion

Roshanaei at al. [1] predicted the mood of the users in Twitter. The users can share their thoughts, feelings, opinions and day to day activities by posting status in the form of images, text. In this work, mood prediction is performed by analysing their day to day postings. According to the mood of the user it will classify the state of the user. The personalised classifiers and positive, negative user's attributes are used. If the user posted a happy status then it is denoted that the user is in happy state and it is classified as positivity. If the user posted a sad status then it is denoted that the user is in sad state and it is classified as negativity. The drawback of this work is that it has no alert system.

Giancristofaro et al. [5] proposed a method to predict sentiment towards transportation in social media using visual and textual features. The dataset in social media is based on Instagram. In this method, a two way communication

between the user and the agency is established by implementing sentiment analysis. The Positives and Negatives about the transportation agency are analysed. Mostly the sentiment analysis is largely based on text processing. The main aim is to utilise the integrated data obtained from the sentiment analysis of both text and images. It uses features of Microsoft Cognitive Services which is web-based to extract high level features of data from the images. It describes the images by detecting the faces and 86 other categories. It uses different machine learning techniques to gain accuracy. Sentiment analysis is useful to find out the positives and negatives of a transportation agency by the users. In this work, the emotions of the users are not recognised.

Uddin et al. [6] proposed a scheme to recognise facial expressions by the salient features using convolutional neural network (CNN). It uses Depth camera based method for facial expression recognition. The LDSP (Local Directional Strength Pattern) concept is used to identify strength pattern. The strength pattern includes low strength and high strength which is similar to positive and negative attributes. The binary values are represented for highest and lowest strength. The six bit pattern is used to recognise facial emotions. There is no alert system. It only predicts the facial expression.

Tasviri et al. [7] presented a model to offer diet to users based on their mood using social network. In this, the user's moods are predicted and they are offered with a diet chart containing healthier food consumption behaviour. The results concluded that based on the user's food consumption, the user's personality and mood are determined.

Cernian et al. [8] developed an application to identify moods and emotions using machine learning techniques. In this work, the mood detector application is designed and implemented to detect the mood and emotional state of the user by analysing 3 parameters of the user such as pulse, skin electro-conductivity and temperature. Machine learning technique is used to analyse the user's mood. This mood detector application is integrated with a music recommender system. It specifically suggests the user to listen to specific playlist based on the currently detected mood. The drawback of this work is that it has no mechanism to alert other users.

Zhu et al. [9] proposed a method to recognise the activities and mood of the users using wearable electronic devices such as smartphones, smart watches etc**.** These wearable electronic devices consist of sensors which read the user's activity and analyse them. If the user is in critical state then the sensor watches will inform the user to take necessary precaution or take medicines.

Taylor at al. [10] predicted the mood, stress, and health using Personalized Multitask Learning. The wearable watches and smart phones are used to determine mood of the user. The ten days data of the user is taken as input to determine the mood of the user the next day.

In [11], depressed mood of the user is predicted based on self-reported histories via Recurrent Neural Networks (RNN). This scheme integrates the categorical embedding layers for predicting depression.

Most of the existing schemes predict the mood of the online user for promoting business activities or for informing the users regarding their schedules. All these schemes do not have an alert system to handle depression. In case of prolonged depression there is no point in warning the user regarding this. It is important to warn the relatives or friends of that user to help him/her get out of prolonged depression.

III. METHODOLOGY

In this section the methodology of the proposed work is presented. This work deals with the mood prediction of users in social media using Deep Learning with Random Forest algorithm (DL-RF). It classifies the users based on the mood and identifies the user under prolonged depression. It alerts

the top five persons in the user's friend list. The system architecture of this scheme is presented in Figure 1.

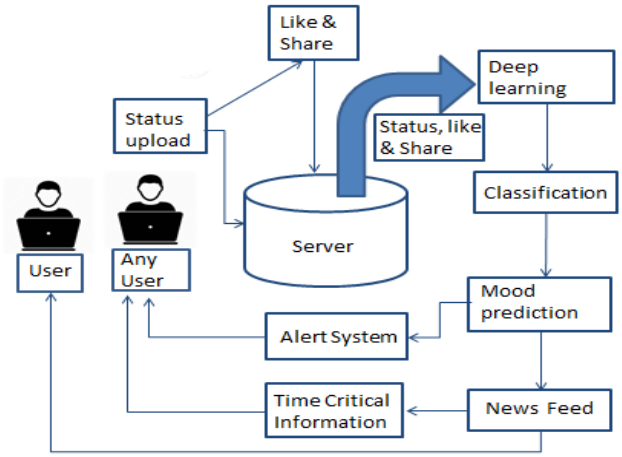

Figure 1. System Architecture of DL-RF Methodology

This scheme involves the following phases:

(i) Input Data Collection
(ii) Deep Learning
(iii) Classification Of Status
(iv) Mood Prediction
(v) News Feed and Alert System.

(i) ***Input data collection***

This phase is to collect the users' status, likes and shares from the social media- Facebook. The average number of images uploaded per day is 300 million as per the statistics [2, 3] and hence there is a large amount of data to be extracted and fed as input to the training phase. In this work about 1000 status posted by users under different mood classification is fed as input to the training phase.

(ii) ***Deep learning***

This phase takes the input dataset and learns to classify and predict the user's mood. The deep learning is an advanced form of training a machine. If the number of training datasets is higher, then the accuracy of the prediction is also higher. The goal of the training phase is to produce accurate results. The features and patterns are extracted from the datasets and analysed. The extracted features and patterns are used to recognize facial expressions and thereby predict the mood of the social media user. The classification of the data is carried out to determine the mood. The evaluation of data is done in order to check the output of the dataset that were not a part of the training datasets. In this deep learning phase, Random Forest algorithm [12] is used along with decision tree algorithm [13] to train the system in classification and prediction. The algorithm for deep learning is presented in Figure 2.

Algorithm: *Deep Learning*

i. *Root nodes are selected based on the best attribute in the dataset.*
ii. *Split the training set into subsets.*
iii. *Repeat the above two steps till the leaf nodes are reached in all branches of the tree.*
iv. *Predict class label for each record in decision tree.*
v. *Compare the values of root attribute with record attribute.*
    a. *Branch corresponding to that value and move to next node until the leaf node is reached.*
    b. *Compare the record attribute value with other internal nodes.*

Figure 2. Algorithm for Deep Learning

(iii) ***Classification of Status***

The user regularly uploads the status in social media. The classification of the status includes happy, sad, angry, contempt, disgust, fear, neutral, surprise. The text or image in the status is analysed and classified based on its positive or negative attributes. When the status is recognised as

disrespectful then it is classified as contempt. When it is recognised as horror, then it is classified as disgust. When there is no reaction (i.e., remaining without expression), then it is classified as a neutral. To classify the status and cluster the classified results, k-nearest neighbours (k-NN) algorithm [14] is used. In k-NN classification, the output is a class membership. This class membership is achieved by calculating the majority of the votes obtained from its neighbours. In the classification phase, k is a user-defined constant, and an unlabelled vector (a query or test point) is classified by assigning the label which is the most frequent among the k training samples nearest to that query point. If the value of k = 1, then the object is simply assigned to the class of that single nearest neighbour. For continuous variables, distance metric such as Euclidean distance is used. For discrete variables, such as for text classification, Hamming distance can be used.

(iv) ***Mood Prediction***

This phase collects the user's status, likes and shares. Based on the status classification, the mood of the social media user is identified by applying random forest algorithm along with tree bagging. Tree bagging [15] averages the values over many samples, and reduces its variance. This helps in predicting the mood of the user especially in prolonged depression over a period of time. Consider a classifier C(S, x), such as a tree. Based on the training data S, it produces a predicted class label at input point x. To bag C, bootstrap samples, each of size N are drawn with replacement from the training data and it is given by equation 1.

$$C_{bag}(x) = \text{Majority Vote } \{C(S^{*b}, x)\} \quad (1)$$

Where b=1 to B.

Bagging can lead to improved prediction accuracy by reducing the variance of unstable procedures (like trees).

(v) ***News Feed and Alert System***

The news feeds are given to the user based on his/her likes. Based on the mood that is predicted by the scheme, the newsfeeds are modified and fed into the user's page. If the user's mood is predicted to be sad or depressed then jokes, funny or motivational facts etc. are presented to the user. These news feeds will help them to get back to normalcy. This helps in preventing suicidal activities. When the user is in the prolonged state of depression, an alert is sent to the user's top five frequent chat friends. This is carried out by the k-NN clustering algorithm [14] which clusters the neighbouring friends in the chat list. The top five frequent users are retrieved using Depth First Search algorithm [16].

IV. RESULTS AND DISCUSSION

The section presents the results of the proposed DL-RF mood prediction scheme. The training model of this scheme involves 1000 status posted by the Facebook users. After this training, the moods of 100 users are classified based on their status posted and it is represented as a pie chart in Figure 3. From Figure 3, it is evident that 36% of the users are in 'happy' mood.

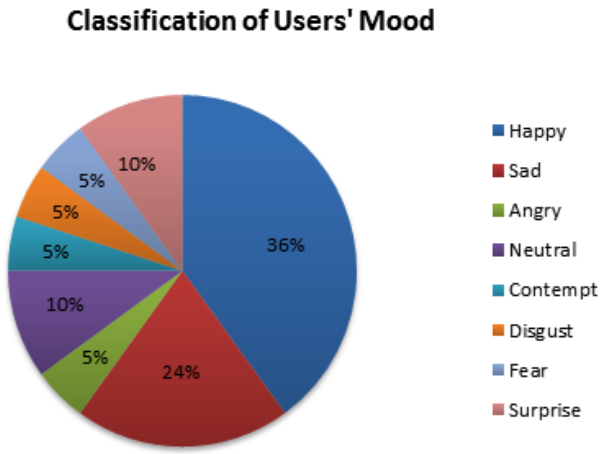

Figure 3. Classification of Social Media Users'Mood

The percentage of users in 'sad' mood is 24% (i.e., 24 users out of 100). Among the 24 sad users, 18 of them are depressed. This implies that 75% of the sad users are depressed. Further analysis, reveals that 14 out 18 depressed users are in prolonged depression. This indicates that 58% of the sad users and 77% of depressed users are in prolonged depression. This is represented in Figure 4.

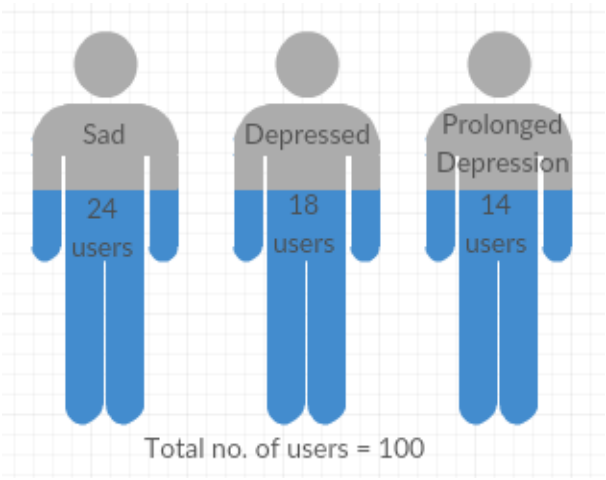

Figure 4. Classification of sad users

Based on the training model, the accuracy of the system varies. As the training dataset increases, the accuracy of mood prediction increases. The work produces about 87% accuracy for mood prediction of social media users.

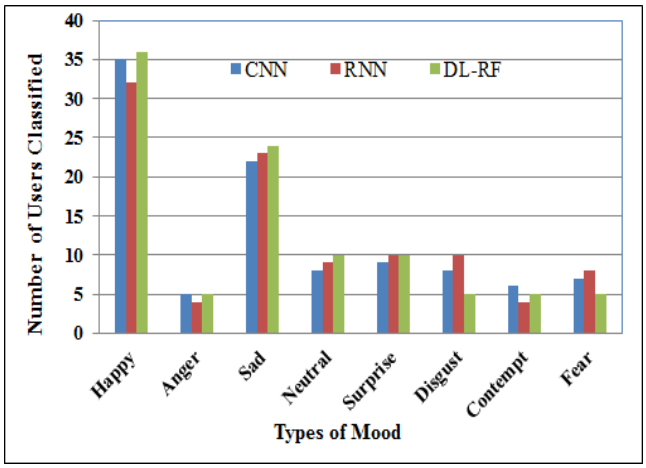

Figure 5.Classification based on different schemes

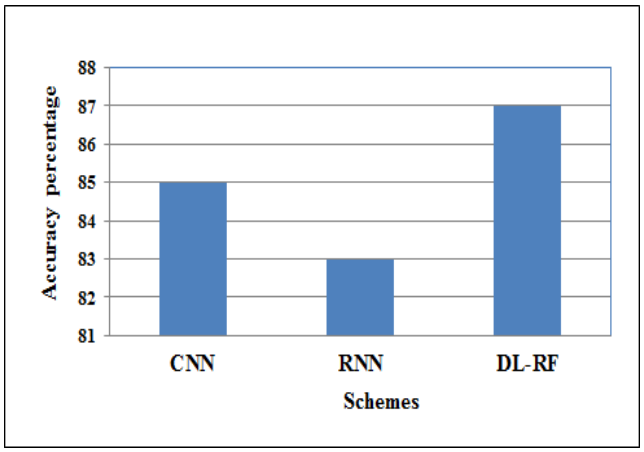

Figure 6. Prediction accuracy of various schemes

The performance of this scheme (DL-RF) is compared with the schemes based on CNN [6] and RNN [11]. The number of social media users classified into different moods based on the various schemes is represented in Figure 5. In Figure 6, the prediction accuracy of these schemes are compared. The prediction accuracy of DL-RF is 87% whereas the prediction accuracy of other two schemes based on CNN and RNN are 85% and 83% respectively. From Figure 6 it is evident that DL-RF outperforms RNN and CNN based schemes.

V. Conclusion and Future Scope

In this paper, the mood of the social media user is predicted in order to identify persons in prolonged depression. It is also used to get back the depressed users in social media to normalcy. Moreover an alert system is included in this scheme to inform the top 5 frequent users in the depressed user's friend list. This scheme produces an accuracy of about 87%. By improving the training data set and the training model, we can achieve better accuracy in mood prediction.

References

[1] M. Roshanaei, R. Han, S. Mishra, "Features for mood prediction in social media", In the proceedings of IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM), pp. 1580-1581, 2015.

[2] A. Hepburn, "Facebook statistics, stats & facts for 2011", Digital Buzz, Accessed 31 January 2018.

[3] D. Noyes, "The top 20 valuable Facebook statistics", Zephoria, Florida, Accessed 10 February 2018.

[4] P. Chiranjeevi, V. Gopalakrishnan, P. Moogi, "Neutral face classification using personalized appearance models for fast and robust emotion detection", IEEE Transactions on Image Processing ,vol. 24, no. 9, pp.2701-2711, 2015.

[5] G. T. Giancristofaro, A. Panangadan, "Predicting Sentiment towards Transportation in Social Media using Visual and Textual Features", In the proceedings of the 19th IEEE International Conference on Intelligent Transportation Systems (ITSC), pp. 2113-2118, 2016.

[6] Md. Z. Uddin, W. Khaksar, J. Torresen, "Facial Expression Recognition Using Salient Features and Convolutional Neural Network", IEEE Access, vol. 5, pp.26146-26161, 2017.

[7] M. Tasviri, S. A. H. Golpayegani, H. Ghavamipoor, "Presenting a Model Based on Social Network Analysis in Order to Offer a Diet to Users Proper to Their Mood", In the proceedings of the 3th International Conference on Web Research (ICWR), pp. 133-139, 2017.

[8] A. Cernian, A. Olteanu, D. Carstoiu, C. Mares "Mood Detector – On Using Machine Learning to Identify moods and Emotions", In the proceedings of the 21st International Conference on Control Systems and Computer Science, pp. 213-216, 2017.

[9] Z. Zhu, H. F. Satizabal, U. Blanke, A. Perez-Uribe, G.Troster "Naturalistic Recognition of Activities and Mood Using Wearable Electronics", IEEE Transactions on Affective Computing, vol. 7, no.3, pp.272-285, 2016.

[10] S.Taylor, E.Nosakhare, A. Sano, R. Picard, "Personalized Multitask Learning for Predicting Tomorrow's Mood, Stress, and Health", IEEE Transactions on Affective Computing, vol. 14, no. 8, 2017.

[11] Y. Suhara, Y. Xu, A. Pentland, "Deepmood: Forecasting depressed mood based on self-reported histories via recurrent neural networks", In the proceedings of the 26th International Conference on World Wide Web. International World Wide Web Conferences Steering Committee, pp. 715–724, 2017.

[12] A. Kaur, R. Malhotra, "Application of random forest in predicting fault-prone classes", In the procedings of the IEEE International Conference on Advanced Computer Theory and Engineering, pp. 37-43, 2008.

[13] A. Gepp, K. Kumar, S. Bhattacharya, "Business failure prediction using decision trees", Journal of forecasting, vol. 29, no. 6, pp. 536-555, 2010.

[14] K. Lee, D. Palsetia, R. Narayanan, M.M.A. Patwary, A. Agrawal, A. Choudhary, "Twitter trending topic classification", In the proceedings of 11th IEEE International Conference on Data Mining Workshops (ICDMW), pp. 251-258, 2011.

[15] L.Breiman, "Bagging predictors", Machine learning, vol.24, no.2, pp.123-140, 1996.

[16] R. Tarjan, "Depth-first search and linear graph algorithms", SIAM journal on computing, vol. 1, no. 2, pp.146-160, 1972.

Authors Profile

Ms.V.Soundarya is pursuing her final year B.E. degree in Computer Science and Engineering at Meenakshi Sundararajan Engineering College, Chennai, India. She is a member of Computer Society of India (CSI). Her main area of interest is Machine Learning. She has done some mini projects based on machine learning.

Ms.R.Swarnalakshmi is pursuing her final year B.E. degree in Computer Science and Engineering at Meenakshi Sundararajan Engineering College, Chennai, India. She is a member of CSI and Indian Society for Technical Education (ISTE). Her areas of interest are compiler design and software engineering.

Ms. I.Swathi is pursuing her final year B.E. degree in Computer Science and Engineering at Meenakshi Sundararajan Engineering College, Chennai, India. She is a member of CSI. Her major areas of interest are machine learning and networking.

Dr.M.K.Sandhya received her Bachelor of Engineering in Computer Science and Engineering from University of Madras and Master of Engineering in Computer Science and Engineering from Anna University. She received her Ph.D under the faculty of Information and Communication Engineering from Anna University in 2014. She is presently working as the Associate Professor in the Department of Computer Science and Engineering in Meenakshi Sundararajan Engineering College, Chennai. She has 12 ½ years of teaching and 6 years of research experience. Her research interests are wireless sensor networks, mobile computing, cryptography and network security. She is a life member of ISTE and CSI.

