Scientific Journal of Impact Factor (SJIF): 5.71

e-ISSN (O): 2348-4470
p-ISSN (P): 2348-6406

International Journal of Advance Engineering and Research Development

Volume 7, Issue 07, July -2020

Depression and Stress Monitoring System via Social Media Data using Deep Learning Framework

Paulami Banerjee[1], Dhanashree A. Phalke[2]

Department of Computer Engineering, D.Y. Patil College of Engineering, SPPU, Pune

Abstract— *Stress and Depression is nowadays are broadly perceived and increasing mental issue that completely influences present society. A health monitoring system that works automatically can be of major importance and also be very critical to improve the sadness and stress recognition framework using social networking site data. The operation of content mining, also the natural language approaches for planning to recognize the feeling or opinion is the function of sentiment analysis. Full of feeling Computing is a way of the examination and advancement of the frameworks and gadgets that can perceive, decipher, process, and mimic the human effects. Sentiment Analysis along with the deep learning techniques could combined together provide us powerful algorithms and frameworks for a given target and observing of mental issues that are, specifically related to depression and stress. In addition, a fundamental plan for incorporating a framework for stress and depression checking is studied. In particular, the paper traces the basic issues and moves comparative with the structure of such a framework.*

Keywords *— Deep learning, Ehealth, stress and depression, sentiment analysis, social media*

I. INTRODUCTION

Social media is surely the richest source of human generated text inputs. Views, feedbacks and critiques which are provided by the social media users will reflect attitudes, thoughts and sentiments towards various topics. There also is a knowledge-based system, that will include an emotional health monitoring system which is used to detect users with a possible psychological disorder specially like depression and stress [1] [4]. Symptoms of all of these mental illness are usually observed in users. In this situation, author argues that online social behaviour extraction offers a great opportunity to find out the mental illness at a very early stage [5].

Stress over a span causes depression a mental disorder, which also has a relevant impact on existing society [5]. Currently, various methods for depression and stress detection are present and also diagnosis rely on self-reporting coupled with a health care practitioners informed assessment. The provision of an effective health monitoring system and also some prediction mechanism aids are important. Sentiment and deep learning technologies could help to tackle these objectives by providing correct methods and ways for objective assessment. Such a mechanism and system do not aim to replace the medical professional but they can support their decisions and aid the same.

The approach is to do emotional health check for emotional problem detection. There are machine learning techniques that does detection of mental problem. Such systems use social data in order to identify with precision possible cases of mental illness [5] [8].

The organization of the paper is as follows: Section II gives the related work, Section III is about the proposed methodology which includes the architecture, algorithm and the mathematical model and last Section concludes the paper with future work followed by references.

II. LITERATURE SURVEY

The literature survey can be divided into two parts as follows:

A. Health Monitoring System using Emotions:

Online Social Sites (OSNs) give important data on different users feeling about various topics. Along these lines, applications, such as, checking and suggestion frameworks (RS) can gather and dissect this information. Here there is exhibited a Knowledge-Based Recommendation System (KBRS) [1], that can incorporate well being observing framework in order to distinguish clients who have potential mental issues, specifically, depression and stress using CNN, BLSTM-RNN algorithms and the eSM2 opinion metric [1]. In cases of microblogs data, various emojis are many of the time utilized and they also have clear meaning in them. They are significant enthusiastic signs for the microblog nostalgic analysis. Such cases have been addressed by developing space as a component portrayal framework and identifying emojis and words into the

passionate space dependent on the semantic composition using enhanced convolutional neural network algorithm [2]. There is also a coordinated website based social networking content investigation stage that uses three level of features, i.e., user produced content like their posts, social graph associations, and user profile information [3], in order to dissect and identify the typical behaviours that go missing altogether from the standard in huge scale social networking sites. Here a few sorts of investigations have been directed for much superior comprehension of the distinctive user practices in the process of discovery of exceptionally versatile indications of the users. This system used PCA algorithm for extracting different features and Profile-Based Collection Technique, Time-Based and Gradual Enhancement Technique for collecting data in real time [3].

Users stress state is firmly identified with that of his friends in web-based social networking, here social associations are considered [4]. Saha and Nguyen gave a technique for the mental health co-occurring in online communities with the help of these features [5]. A target building prescient model that influence language and standards of conduct, these also are used to decide if a user is experiencing the instances of mental issue. These prescient models are made conceivable by utilizing a novel information assortment process, authored as Subconscious Crowdsourcing, which gathers a quicker and progressively solid dataset of patients. Few tests recommend that extricating explicit language examples and social connection highlights from solid patient datasets can enormously add to advance examination and identification of mental issue [6]. As an alternative approach one can provide with evidence that daily stress can be correctly identified based on behavioural metrics, derived from a user's activity on his mobile phone and also from additional indicator, like the given weather and the personality traits [7]. The multifactorial statistical model, which is person-independent used TF-IDF and LIWC algorithms for feature Extraction and Random forest for classification [7]. As Schuller and Batliner found the SVM algorithm is also used a lot for emotion classification tasks [12]. Also as per Alghowinem, SVM is used in the depression detection using classification concept too [13].

Understanding action between users in the Facebook social network in order to catch the idea of mental health awareness can be done [8]. It is observed that if we consider the activities they have different links associated with them. These have the tendency to form and go in a very quick manner in the reference to time frames. The strength of the ties however have a tendency to in general decrease trend for the given activity as the link of the social network has spent more time [8]. Social media tools devices are wide spread in web correspondence and are picking up prominence in the correspondence procedure between open organizations and residents [9]. As per Berbano, apart from social media data there are also other ways to identify user's stress levels. There work on the neural engineering to focus on the classification of various stress such as emotional stress or physical stress or mental stress or no stress with Electroencephalography (EEG) signal analysis [10]. So, stress may be classified through the use of the EEG signal analysis which may be acting as the objective functions for the same. The features of the current EEG recordings are next sent for the step of pre-processed and they are later selected with the help of Discrete Wavelet Transform. These features are then used as inputs to classify stress using different Neural Networks and validated using K-times Cross Validation Method [10].

B. Sentiment Analysis:

Sentiment analysis can help us in framing a stress monitoring system [11]. Sentiment analysis proposed by S. Rendle gave methods using machine learning [16]; and also using lexicons. [14] [15]. An effective half mixed type model - a factor diagram model joined with Convolution Neural Network to use the content or data of the tweet and social collaboration data of the user for stress discovery can be used for this [4]. There is also a coordinated website related social networking content investigation stage that uses three level of features, i.e., user produced content like their posts, social graph associations, and user profile information [3]. As Araque pointed out, Sentiment features needed to be extracted, there can be so many kinds of ways and deep learning can also serve the purpose [17]. Also, neural networks have their performance improved if they are combined with manual extraction of the features [17]. Also various ensemble techniques may be combined together and the resultant accuracy also improves for the same. With the use of this the ultimate goal is to identify the views type that is if user is feeling positive, negative [18].

III. **PROPOSED METHODOLOGY**

In the proposed approach, the given task may be formulated to behave like a classification problem. For the detection of the different emotional or the psychological issue for the OSN user with the help of social network data. It uses sentiment analysis and also the deep learning methods (specifically RNN) for finding:

i. Stress
ii. Depression

iii. Positive comments
iv. Negative comments

It is an solution that aims to detect those potential users who may be facing any emotional disorder, according to the classification of sentences with depressed or stressed content. Also the comments of the users are analysed and their positive or negative mind frame may be identified in this process. The users are each identified and also they are sent with motivational messages. This helps the users to boost up their mood.

A. Architecture:

The system is composed of data extraction, pre-processing, classification, sentiment analysis and motivational message sending blocks. Below is the architecture diagram of the system:

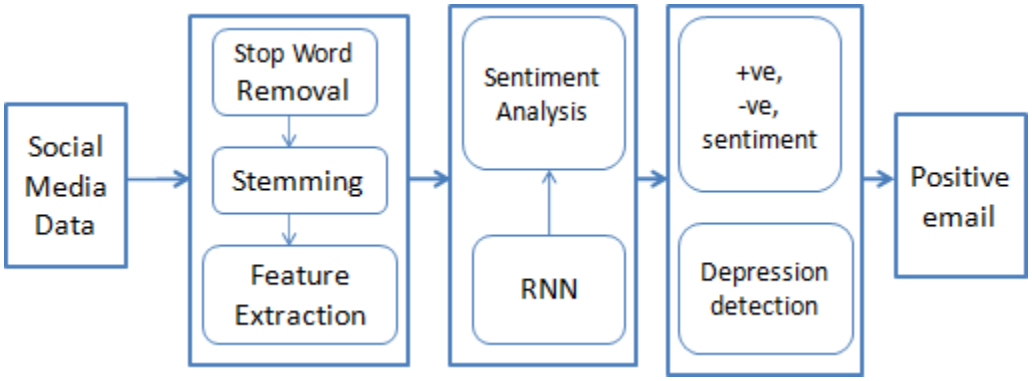

Fig. 1. Proposed System Architecture

In data collection phase, Twitter and Facebook realtime data via API is used. Next step is of performing data pre-processing where first Stop word Removal is carried out. This is done using Natural language toolkit's function 'stopwords'. This technique removes stop words like: is, are,they,but. Next stemming is performed, it can be done with 'PorterStemmer' of Natural language toolkit. It removes the suffix and prefix of a word. Output of this process is the word in its original form. For example, played becomes play and clustering becomes cluster. Features are extracted next using TF-IDF these are then given to RNN for classification and sentiment analysis is carried out with the help of Naive Bayes.

B. Algorithm:

The main algorithm used here is as follows:

Recurrent Neural Network:
"There may be cases like when it is required to predict the next word of a sentence, the existing previous words are used for reference. For such cases RNN came into existence, they solved this issue with the help of a Hidden State. The main and most important feature of RNN is Hidden state, it will remember information about a sequence [19]. Hence we will obtain a better classification algorithm with the help of previous data too.

Steps:
Let us consider a network with one input layer, hidden layers and one output layer.

- To the network give a single step input.
- Next calculate the current state using set of current input and the previous state.
- After all the time steps are completed the final current state will be used to calculate the output.
- The output is then compared to the target output and the error is generated.
- The error will be next back-propagated to the network to update the weights and hence the network (RNN)
- will be completing its training and it will be used for classification. "

C. Mathematical Model:

The mathematical model for the system is as-

S = {I, F, O}
where,
I = Set of inputs
The input consists of set of Words/Emoticon. It uses Twitter/Facebook dataset.

F = Set of functions where F = <F1, F2>
F1: Data Collection
Dataset =< T, F >
where, T- Twitter and F-Facebook
F2: Classification
Data =< w, rnn >
where, W – Words and rnn – Recurrent Neural Network

O = Output (i.e. Positive comments, Negative comments, Stressed user, Depressed user)

IV. **RESULTS AND DISCUSSION**

A. Dataset:

The dataset in training phase is build to classify depressed, and non-depression expressions, was built using sentences written by users on an OSN. Data Set from kaggle are taken and are combined together.

• Sentiment 140 [20]
• Depression Data Set [21]

The dataset is split into 80 percentage for training and 20 percentage for testing. The most frequently occurring words in the data can also be identified and we can plot them as per most number of occurrence. The data is provided with label ’0’ and label ’1’ accordingly. Post training testing is carried out and we identify the models performance.

B. Results:

The results of the system is based on the following metrics: accuracy, precision, recall and F1-Score. Accuracy is sum of correctly classified sentences over total number of input instances:

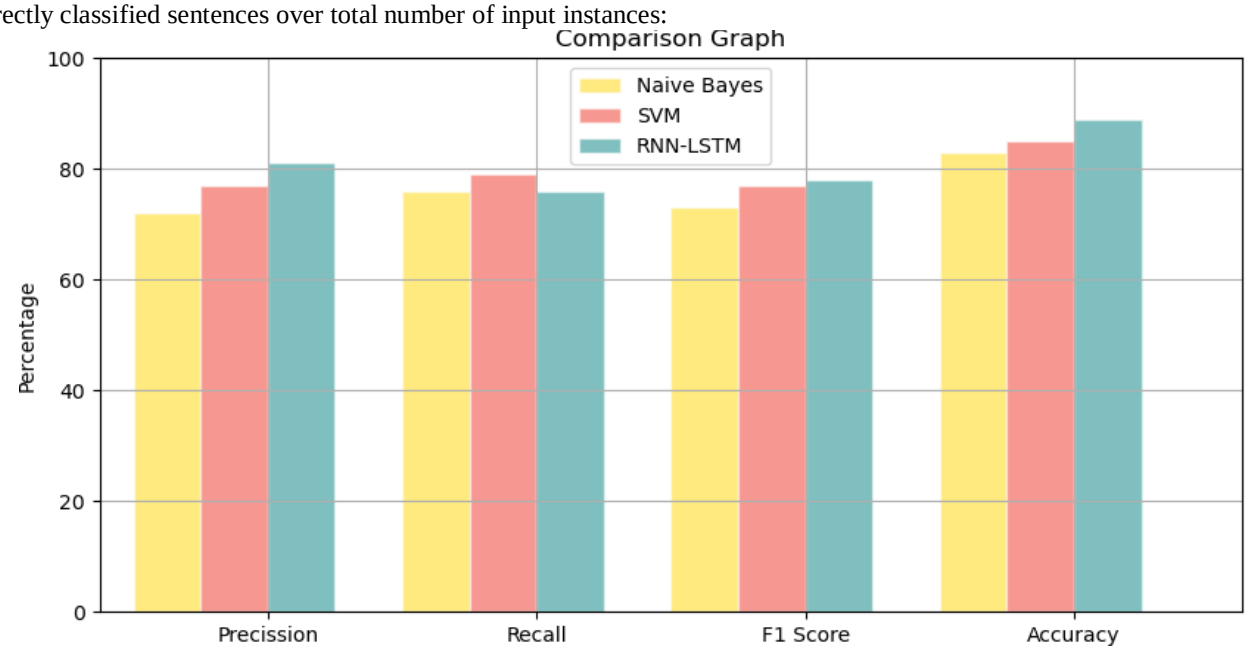

Fig. 2. Performance Analysis Graph

On this basis, we can calculate four measurements:
Accuracy = TP+TN/TP+FP+TN+FN
Precision = TP /TP+FP
Recall= TP/TP+FN
F1-Measure = 2*(Recall * Precision) / (Recall + Precision)

Table 1. Comparison Table

	Naive Bayes	SVM	RNN-LSTM
Precision	77.33	77.48	81.53
Recall	76.67	79.16	76.64
F1-Score	73.94	77.98	78.34
Accuracy	83.72	85.13	89.78

V. **CONCLUSION**

With this system, one can find out the site users who might be dealing with depression and stress. Such issues are threatening, they impact people's health in today's society. Thus users suffering from depression can be identified and they can be helped before they end up taking any drastic step which might have a very long lasting impact. Using the data of social networks for the real world as a basis, study is done on the relation between the psychological disorders and also provide health precaution by sending them with motivational messages. The accuracy of RNN-LSTM was around 0.89 to identify depression. In future, this can be made part of a bigger system or it may be merged to a new concept, or other platforms containing images may be explored.

REFERENCES

[1] Renata L. Rosa, Gisele M. Schwartz, Wilson V. Ruggiero, and Demostenes Z. Rodrıguez, Senior Member, IEEE" A Knowledge-Based Recommendation System that includes Sentiment Analysis and Deep Learning" IEEE 2019.

[2] Guang Yang, Haibo He, Fellow, IEEE, and Qian Chen" Emotion-Semantic Enhanced Neural Network" IEEE 2019.

[3] 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, Feb 2018.

[4] 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, Sept 2017.

[5] Budhaditya Saha, Thin Nguyen, Dinh Phung, Svetha Venkatesh" A Framework for Classifying Online Mental Health Related Communities with an Interest in Depression" IEEE 2016.

[6] Chun-Hao Chang, Elvis Saravia, Yi-Shin Chen" Subconscious Crowdsourcing: A Feasible Data Collection Mechanism for Mental Disorder Detection on Social Media" 2016 IEEE/ACM

[7] Andrey Bogomolov, Bruno Lepri, Michela Ferron, Fabio Pianesi, Alex (Sandy) Pentland," Daily Stress Recognition from Mobile Phone Data, Weather Conditions and Individual Traits" IEEE Conference 2015

[8] Bimal Viswanath Alan Mislove Meeyoung Cha Krishna P. Gummadi," On the Evolution of User Interaction in Facebook" ACM 2011

[9] I.-R. Glavan, A. Mirica, and B. Firtescu, "The use of social media for communication." Official Statistics at European Level. Romanian Statistical Review, Dec. 2016.

[10] 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," MDPI, Sept 2017.

[11] R. Rosa, D. Rodr, G. Schwartz, I. de Campos Ribeiro, G. Bressanet al., "Monitoring system for potential users with depression using sentiment analysis," in 2016 IEEE International Conference on Consumer Electronics (ICCE). Sao Paulo, Brazil: IEEE, Jan 2016.

[12] B. Schuller, A. Batliner, S. Steidl, and D. Seppi, "Recognising realistic emotions and affect in speech: State of the art and lessons learnt from the first challenge," Speech Commun., Nov. 2011.

[13] S. Alghowinem, R. Goecke, M. Wagner, J. Epps, M. Hyett, G. Parker, and M. Breakspear, "Multimodal depression detection: fusion analysis of paralinguistic, head pose and eye gaze behaviors," IEEE Transactions on Affective Computing, 2017.

[14] R. L. Rosa, D. Z. Rodrıguez, and G. Bressan, "Music recommendation system based on user's sentiments extracted from social networks," IEEE International Conference on Consumer Electronics (ICCE), Jan 2015.

[15] R. L. Rosa, D. Z. Rodrıguez, and G. Bressan, "Sentimeter-br: A new social web analysis metric to discover consumers' sentiment," IEEE International Symposium on Consumer Electronics (ISCE), Jun 2013.

[16] S. Rendle, "Factorization machines with libfm," ACM Trans. Intell. Syst. Technol., May 2012.

[17] O. Araque, I. Corcuera-Platas, J. F. Sanchez-Rada, and C. A. Iglesias, "Enhancing deep learning sentiment analysis with ensemble techniques in social applications," Expert Systems with Applications, 2017.

[18] Xing Fand and Justin Zhan, "sentiment analysis using product review data", Springer, 2015.

[19] Don Hush, Chaouki Abdallah and Bill Horne, "High Order Recursive Neural Networks," University of New Mexico,1991.

[20] https://www.kaggle.com/kazanova/sentiment140.

[21] https://www.kaggle.com/josechipanatica/depression-twitter.

