A Machine Learning Approach for Detecting Stress Based on Social Interactions in Social Networks

Gokila Arthi.S[1], Mrs.S.Sandhya[2]
[1]M.Sc Computer Science, Department of Computer Science, Sri Krishna Arts & Science College, Coimbatore, India
[2]Assistant Professor, Department of Computer Science, Sri Krishna Arts & Science College, Coimbatore, India

Abstract- *Stress is a kind of the demand in your body's way of responding to any. It can be based on Both Good and bad experiences. Psychological stress is threatening people's health. With the reputation of a social media network, people are used to sharing their schedule and daily activities and interacting with friends on social media platforms, making It is feasible to leveraging the online social network data for stress detection. Data mining techniques are used for a various number of applications. In industry, data mining plays an important role in Detecting Stress In this paper, propose a new model to detect stress. In this model, initially, find a correlation of users stress states and social interactions effectively. This defines set of stress-related textual, visual, and social attributes from various aspects and proposes a novel hybrid model combined with Convolutional Neural Network CNN to leverage tweet content and social interaction information for stress detection effectively. From the experimental results, the proposed model can improve the detection performance reaches 97.8% accuracy which is faster than the existing system.*

Keywords- Data Mining, Machine Learning, Stress detection, Stress Detection, social media, social interaction.

I. INTRODUCTION

Stress is called simply feeling of being overloaded, wound uptight, tense and worried. It can sometimes help to motivate us to get a task finished or completed or perform well. But the stress it can also be harmful if we become over-stressed and it interferes with our ability to get on with our normal life for too long. Generally, Psychological Stress is one of the mental stresses and it is becoming a threat to People's Health Nowadays. In the rapid pace of life is mostly filling the feeling of stress, more and more people. According to a worldwide survey reported over half of the population has experienced an appreciable rise in stress over the last two year. In this Stress, results are analyzed from the tension between an individual's reaction to difficulties or challenges and his or her ability to handle and resolve the stressful situation. How many numbers of peoples cope with stress depends on the resources? So that is available to them and whether they have the skills to utilize these resources. Generally, stressor relates to challenging the issue that may produce the stress. Acute stress reflects short-lived stress or – event full experiences that occur once or multiple times. The Stress Example of acute stress is can result from relocating to a new town or having a serious disagreement with a parent. The resultant of experiences is negative or positive impacts on health and its development depends on how one perceives and copes with the stressful experience. Acute stress that results from an unforeseen external event or significant life of the changes such as the death of a parent, changing homes or schools, or being in a car accident. This type of stress changing is upsetting particularly because the individual experiencing this stress doesn't have control over the event. An ongoing form of stress is called chronic stress that occurs as a part of one's daily life and that continually taxes one's physical and mental resources. The Typical stress categories of chronic stressors that children and youth experience include frequent parental arguments, chronic illness, neighborhood crime, caregiving for a parent or sibling and trying to adapt to another culture. When a child or adolescent is experiencing difficulties coping with stress, he or she is likely to show changes in mood, behavior, and/or physical appearance in SIGNS OF STRESS. Physical changes are based on some reasonable changes such as lacking in sleep, muscle tension, headache, stomachache, lacking intake of food and lack of energy in the human body. Emotional changes are based on heart feelings such as nervousness, anxiety, loss of enthusiasm about things he or she used to enjoy, anger, shyness, and feelings of helplessness and hopelessness. Behavioral changes are based on good and bad timings or situations that include poor eating habits and excessive weight gain/loss over a short period of time. Being able to recognize the signs of stress is an important skill for children and youngsters.

II. PROBLEM DEFINITION

In [1] Jia Jia presents the concept namely, 'stress detection method automatically using cross-media microblog data'. In this approach set of low-level features is obtained from the tweets. Then, psychological and art theories are defined and extracted: linguistic attributes from texts, visual attributes from images, and social attributes from comments are based on middle-level representations. The stress categories may learn to create by the Deep Sparse Neural

Network (DSNN). In this proposed method is effectively detects the stress of psychological and from a microblog of data. It is a very efficient one for stress detection.

In [2] Ling Feng presents a deep neural network (DNN) model. In this approach investigated the correlations between stress state of the user and their tweeting content, social engagement and behavior patterns. After that stress-related attributes are defined as follows: 1) low-level content attributes from a single tweet, including text, images, and social interactions; 2) user-scope statistical attributes through their weekly micro-blog postings, tweeting time collecting the leveraging information, tweeting types and linguistic styles. The combination of content and statistical attributes, by a convolutional neural network (CNN) with cross autoencoders, to generate user-scope content attributes. Finally, a DNN model is proposed to incorporate the two types of user-scope attributes to detect users' psychological stress. The proposed model provides an effective experimental result. It describes how to detecting micro-blog of data creating psychological stress.

In [3] Jichang Zhaol presents the concept namely, "Mood Lens". Generally, 95 percentage od emoticons are mapped into four categories of sentiments such as angry, disgusting, joyful, and sad. These sentiments act as the class labels of tweets. The Training data is collected and it is around 3.5 million labeled tweets and after a Naive Bayes classifier is trained, with an empirical precision of 64.3%. Using Mood Lens for real-time tweets obtained from Weibo, several interesting temporal and spatial patterns are observed. Also, sentiment variations are well-captured by Mood Lens to effectively detect abnormal events in China.

In [4] H. Lin, J. Jia, Q. Guo, L.Feng, Y.Xue, L. cai, J.Huang, presents the concept namely, "Psychological stress detection from cross-media microblog data using deep sparse neural network". In this concept build up a three-level system to characterize the issue. To start with getting to set of low-level highlights from the tweets. At that point characterize and extricate center level outlines in light of mental and craftsmanship plans: phonetic characters from tweets' writings, visual characters from tweets' pictures, and social characters from tweets 'remarks, retweets, and top choices. At long last, a Deep Sparse Neural Network is made to take in the pressure classes coordinating the cross-media characters. The trial comes about the view that the proposed strategy is powerful and effective in recognizing mental worry from microblog information. The future work is to research the social connections in mental worry to additionally enhance the location execution.

In [5] Yuan Zhang, Jie Tang, Jimeng Sun, Yiran Chen, presents the concept namely"Emotion prediction via dynamic continuous factor graph model". In this concept present MoodCast process based on a dynamic continuous factor graph model for modeling and conclude users' emotions in a social network. Mood Cast formalizes the problem into a dynamic continuous factor graph model and defines three types of factor functions to capture the different types of information in the social network. For model learning, it uses a Metropolis-Hastings algorithm to obtain an approximate solution. Experimental results on two different real social networks demonstrate that the proposed approach can effectively model each user's emotion status and the prediction performance is better than several baseline methods for emotion prediction.

In this paper [6] Lexing Xie and Xuming, presents the concept namely "Picture tags and world knowledge: learning tag relations from visual semantic sources". Proposes novel methods to analyze photo tags and tag relationships, using data from Flickr, Image Net and Concept Net. A novel network inference algorithm, ICR is designed to estimate latent relationships from tag co-occurrence. This method obtains tag statistics on thousands of tags from millions of images. The proposed tagging algorithm generalizes to unseen tags and is further improved upon incorporating tag-relation features obtained via ICR. The future work, of this paper, is techniques to better incorporate multi-word terms and out-of-vocabulary words; advanced NLP techniques for learning word relations from the free-form text; evaluation of latent concept relation suggestion, and predicting the type of relations.

This paper [7] Author Chi Wang, Jie Tang, Jimeng Sun, proposes a pairwise factor graph (PFG) model to model the social influence in social networks. In this paper formulate and tackle the problem of dynamic social influence analysis and present a pairwise factor graph (PFG) model to model the pairwise influence. Specifically, the influence between two users is modeled as a marginal probability of two hidden variables in the factor graph model. An efficient learning algorithm is proposed. Next, propose a time-dependent factor graph (DFG) model to further incorporate the time information, which is described as a factor function across time windows. Thus influence is propagated across social networks of different time windows. Experimental results on three different types of data sets show that the proposed approach can effectively discover the dynamic social influences. In future work apply the inferred social influence to help influence maximization. Parallelization algorithm is used.

In [8] J. B. Yang, M. N. Nguyen, P. P. San, X. L. Li, presents the concept namely "Deep Convolutional Neural Networks On Multichannel Time Series For Human Activity Recognition". In this paper propose a systematic feature learning method. This method adopts deep convolution neural networks (CNN) to automate feature learning from the raw inputs in a systematic way. The learned features are deemed as the higher level abstract representation of low-level raw time series signals via the deep architecture. In the supervised learning is leveraging the labeled information, the learned features are endowed with more discriminative power. The Unified model is mutually enhancing the feature learning and classification. The benefit-based this unique advantages of the CNN make it outperform other HAR algorithms, as verified in the experiments on the Opportunity Activity Recognition Challenge and other benchmark datasets.

III. PROPOSED SYSTEM

The chapter discusses the proposed methodology and the steps involved in this proposed system. The previous framework is requires more training dataset and time, so it is drawbacks of existing system. The large data sets are computationally demanding. Stress detection performance is low. Users have affected high psychological stress and it may exhibit low activeness on social networks. To address this problem proposed system proposes a new effective Framework called Improved Learning and Inference by Factor Graph (ILIFG). This framework aims at increasing the classification performance and detection performance through the hybrid approach. The goal of the proposed system is applying the Multi-Class learning process to identify the best class of objects and classifying them accordingly. This aims at producing the least false alarm rate and improving the classification performance.

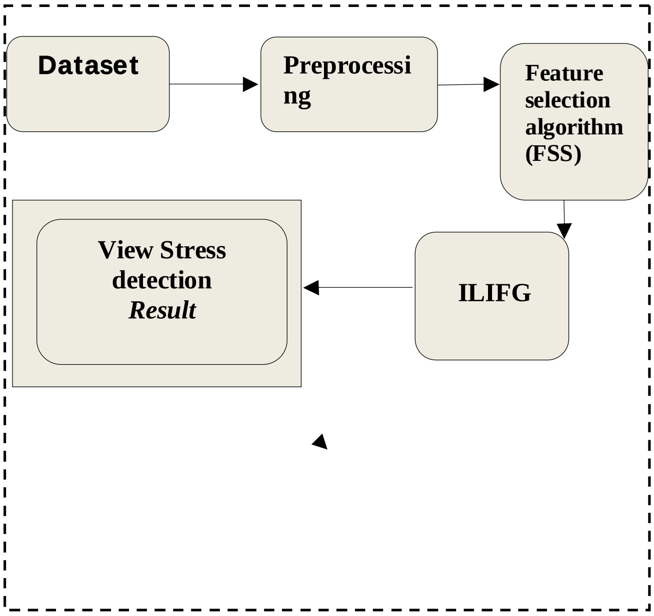

Figure 3.1 *Proposed system architecture*

A. Contributions of the proposed System

The followings are the contributions of the proposed system.

- The system implements a new improvised Learning and Inference by Factor Graph (ILIFG) Model for improves the stress detection performance.
- This also creates a new advanced latent factor model to reconstruct the missing data for fast and accurate stress detection. The system developed with the intention of high accuracy and less training overhead.

IV. METHODOLOGIES

A. Latent factor model

Data preprocessing is a data mining real world data is often incomplete, inconsistent, or lacking in certain behaviors or trends, and is likely to contain many errors in that data. Data preprocessing is one of the proven methods of resolving such issues, problems and so on. Data preprocessing prepare raw data for further processing. Proposed used latent factor model to reconstruct the missing data from the records collected from a UCI in a different unit.

B. Improved Learning and Inference by Factor Graph (ILIFG)

This is a type of probabilistic graphical model. A predictive model analysis in data mining is a procedure by which a model is generated or selected to predict the best likelihood of an outcome. In certain scenarios, the model is selected on the basis of detection theory to guess the probability of an outcome given with a group of input data.

Input: a series of time-varying attribute augmented network G with stress states on some of the user nodes, learning rate h;

Output: parameter value full stress state vector Y;

Step1: Import the libraries
Step2: Import the data-set
Step3: Check out the missing values
Step4: process continues until total dataset length
Step5: randomly initializes Y;
Step6: Initialize model parameters u; Repeat
Step7: Compute gradient;
Step8: Update full stress state
Step9: Update probabilistic graphical model

V. RESULT AND ANALYSIS

Metrics	**Dataset**	**Existing**	**Proposed System ILIFG**
	DS1(50)	95	99
Detection Accuracy (%)	**DS2(100)**	93	98.8
	DS3(120)	93	98.5
	DS4(150)	90	98

A. Experimental Results

This section describes the implementation process. Implementation is the realization of an application, or execution of plan, idea, model, the design of research. This section explains the software, datasets, and modules which are used to develop the research. Then experimental term is performed on an Intel I3 Processor with a RAM capacity 4GB.

B. Data Set

The data used in this study contains real-time Twitter data and the data store in the database. The dataset is generally composed of structured and unstructured text data. This data contains several attributes such as users' tweets. In this user, tweets contain much more words from categories such as death, sadness, anxiety, anger, and negative emotion, while non-stressed users' tweets contain more words from categories like friends, family, affection, leisure, and positive emotion. There are a total of more than 500 tweets records in the database.

C. Results and Analysis

The experiments are designed so that the different parts of the work could be evaluated. These include the evaluation of the features of the above dataset, the feature selection and also the feature creation methods. Finally, the performance of this proposed work Scheme was compared with the existing algorithms based on the following parameters.

- **Accuracy** – Determines the correctness
- **Precision** –Repeated process same result
- **Time taken** – Determines the processing time involved.

The dataset is preprocessed by the latent factor model and features are selected effectively and finally, the prediction process is made by ILIFG Model.

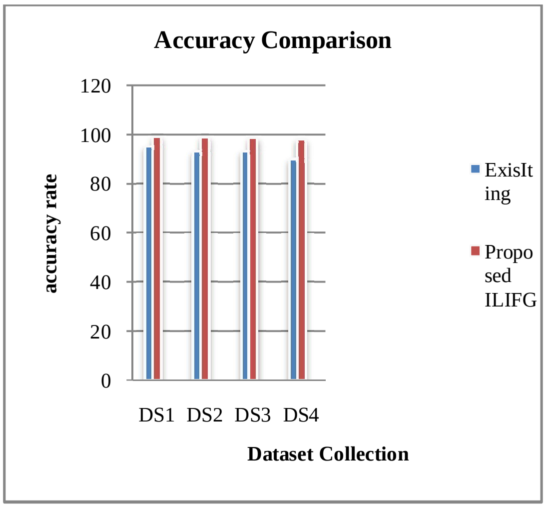

***Figure 5*.1** Accuracy Comparison

D. Performance comparison of the proposed system with existing approaches based On Stress prediction Result accuracy

From the results shown in the graphs, it can be observed that the proposed model-based approaches provide better accuracy and increased true positive rate when it is analyzed with the different number of datasets. The system finally performs the analysis to show the accuracy of the proposed system in graph format.

V. CONCLUSION

The study and research proposed a new prediction scheme for stress detection. The system studied the main two problems in the literature survey, which are prediction accuracy and classification delay. The study overcomes the above two problems by applying the effective enhanced ILIFG Model. The system effectively identifies and prediction the stress. The experimental result shows that the integrated extended proposed algorithm shows better quality assessment compared to traditional research techniques. From the experimental results, the prediction accuracy of our proposed algorithm reaches 98.8% with a convergence speed which is faster than the existing system. As further work, improvements can easily be done since the coding is mainly structured or modular in nature. In the system can change the existing modules or adding new modules can append improvements. Further enhancements can be made to the application by expanding the existing modules future research may use the model to identify the existing area of research in the field of data mining in other dataset and use of other classification algorithms. As further work, use this model as a functional base to develop an appropriate data mining system for classification performance.

REFERENCES

[1] H. Lin, J. Jia, Q. Guo, L.Feng,Y.Xue, L. cai,J.Huang, Psychological stress detection from cross-media microblog data using the deep sparse neural network, In April 2014.

[2] Yuan Zhang, Jie Tang, Jimeng Sun, Yiran Chen, and Jinghai Rao.Moodcast: Emotion prediction via dynamic continuous factor graph model, IEEE International Conference on Data Mining 2016.

[3] Lexing Xie and Xuming He. Picture tags and world knowledge: learning tag relations from visual semantic sources, In ACM October 2018.

[4] Chi Wang, Jie Tang, Jimeng Sun, and Jiawei Han Dynamic social influence analysis through time-dependent factor graphs, IEEE 2016.