

A Study on Stress Prediction using Sentiment Analysis on Social Media

Dr. P. Suganya¹, P. Bharathi², R. Rinipriya³

¹ Assistant Professor, Department of Computer Science, Erode Arts and Science College (Autonomous) Erode, Tamilnadu, India, Email: suganya.rithikjos83@gmail.com

² Assistant Professor, Department of Computer Science, Erode Arts and Science College (Autonomous) Erode, Tamilnadu, India, Email: bharathimca007@gmail.com

³ Assistant Professor, Department of Computer Science with Data Analytics, St. Joseph's College for Women Tirupur, Tamilnadu, India, Email: rini.priya7@gmail.com

Abstract:

Stress is an important factor in a person's daily life. Sentiment analysis is the study of people's emotions, thoughts, and feelings. According to the World Health Organization (WHO), stress is a mental health issue that affects the well-being of individuals worldwide. In today's world, people experience different levels of stress, which can influence emotions, thoughts, and behavior. Therefore, it is essential to identify and manage stress before it becomes a serious problem.

A large amount of informal data is shared daily on social networking sites, blogs, and discussion forums. This paper presents an approach for detecting stress using information collected from social media platforms. Social network data can be utilized effectively to identify stress levels among individuals. Various methods are available for stress detection, including medical tests and analytical techniques. This paper discusses different methods used in sentiment analysis and classifies them based on their advantages and limitations for the stress detection process.

Keywords: Social media, Sentimental analysis, Stress level Protection, Mental health, Support Vector machine (SVM), Machine Learning.

1. Introduction

Today modern world that create the relationship is being impacted by the digital revolution. Modern technology diverts time and attention from friends and families. As grow of rise in interaction of people in social media platforms, making it feasible to identify online social network data for stress detection. [6] People they are affect from stress is constructive because it makes them stronger to keep active in their work and it motivates them, and also stress is destructive in another few people because it makes them feel alone, depressed, inactive, mood swings, and it can affect their mental health. Stress problems have been divided into two parts one is a short-term problem like headache, sweating etc., and another one is a long-term problem like changes eating habit in persons, loosing body weight etc., [12]. With the development of social networks like Twitter more and more people are willing to share their daily events and moods, and interact with friends through the social networks. As these social media data timely reflect users' real-life states and emotions in a timely manner, it offers new opportunities for representing, measuring, modelling, and mining user's behaviour patterns through the large-scale social networks, and such social information can find its theoretical basis in psychology research. The paper focuses on various stress detection methods to find the user's psychological stress state, and able to identify the higher efficiency stress detection method. Sentimental analysis is the determinative the emotion tone behind the word. That understands of the opinion, attitude, and emotions through the word. The aim to utilize machine learning technique for stress detection on social media.

2. Survey On Stress Detection

Preethi Rajam C.R., Uma Maheswari N., at el., (2019) the social media analysis is the process of finding and analyzing the user's thoughts and opinions which also one of the important aspects in the developing business environments. The machine learning and deep learning methods to perform the classification of stress analysis. Here both image and text- tweet data are used and the images are processed with the Optical Character Recognition and the text data are processed by using the Natural Language Processing and Convolutional Neural Network for classifying the tweet content of the user as stressed or non-stressed.[1]

Akshada Kene, Dr. Shubhada Thakare, et al., (2021) Stress is a major symptom for mental health. Stress affects every process of a person's life such as feelings, emotions, thoughts, and behaviours. The proposed system for stress detection of users based on the Support Vector Machine (SVM) and Random Forest (RF) algorithm. In this review, to analyze how various factors affect the stress of users. Performance evaluation was measured using accuracy, precision, and recall parameters. It was observed that SVM gives better accuracy (80.2%) than RF. [2]

Reshma Radheshamjee Baheti, Supriya Kinariwala et al., (2019) this review presents a method to find stress level and relaxation on twitter dataset i.e. working on sentiment analysis to find thoughts or behaviours about daily life. Sentiment analysis works the automatic extraction of sentiment related information from text. Here using Tensi Strength framework for sentiment strength detection on social networking sites to extract sentiment strength from the informal English text. Here, SVM with Ngram gives better result Precision is 65% and Recall is 67%. But, the main objective of this technique is to find the explicit and implicit amounts of stress and relaxation expressed in tweets. [3]

S. M. Chaware, Chaitanya Makashir, Chinmayi Athavale et al., (2020) the rapid increase of mental problems or depression has become a one of the challenge to human health and life. It is difficult to timely detect mental disorders or stress for proactive care. machine learning techniques including the use of various algorithms such as Decision Tree, Naïve Bayes, Random Forest, etc. which gives a lower accuracy of 70% on average. In this paper, we are using a closeness of stress levels with social media data shared by many users. The TSVM algorithm best classifies data than machine learning algorithms like Naïve Bayes, Random Forest, Decision Tree. [4]

Mummadi Rama Chandra, Boda Nithin Kumar et al., (2023) Stress is a very universal human experience Especially in the online world. The proposed system helps predict stress people, thereby helping society in solving serious problems. Proposed a system used Bernoulli Naïve Bayes algorithm to detect users mental stress by leveraging users tweet content and social interactions. And the result would be 74% accuracy with BernoulliNB. [5]

P.Iswarya, D.Maheswari et al., (2020) the stress detection methods such as Support vector Machine, Probabilistic Neural Network and Random forest are modeled and incorporated with time variant tweet content attributes and social interactions used for learning and detecting user emotion states. To fully leverage both content and social interaction information of users' tweets, we find a efficient model which combines the CNN with a SVM, RF and PNN. In this review, several emotions of the user are predicted and more negative emotions reveals psychic stress to that student. The RF model outperforms 5-6% higher than other two classification methods. [6]

S.Venkateswaran, K.Sangeetha et al., (2018) A set of stress-related textual, visual, and social attributes from various aspects are first defined and then propose a novel hybrid model. K-mean clustering techniques in order to cluster the user data and provide accuracy for the user stress levels.that are gathered and provided by graph internet explorer. In this system, to display a system for distinguishing users psychological stress states from clients. [7]

Sameer Dev Sharma, Sonal Sharma et al., (2022) Predicting the stress levels of working professionals is one of the most time-consuming and difficult research topics of current day. The model proposed here comprises dataset preparation, feature extraction, optimal feature selection, and classification using DRNNs. The system's main advantages are its simple design structure, efficient performance, accurate classification, and reduced training model time. The proposed technique's performance is validated and compared to other state-of-the-art models using various measures in experiments. The proposed DRNN technique outperforms the other techniques in terms of accuracy, sensitivity, specificity, precision, recall, and error rate, as evidenced by the results. [8]

Moein Razavi, Samira Ziyadidegan et al., (2022) to explore the field of ML in mental health, to used a scoping review process. This literature identified application domains of detection and prediction of stress as a

contributor to mental health disorders. According to this review the best AI algorithms for stress detection were found to be Random Forest and Neural Network. [9]

Thilagavathi. P, Pushkala. K at el., (2018) Sentiment Analysis (SA) is an ongoing field of research in text mining field. To present a framework for detecting users psychological stress states from users' weekly social media data, leveraging tweets' content as well as users' social interactions. Employing real-world social media data as the basis, we studied the correlation between user' psychological stress states and their social interaction behaviors. In this review discovered several intriguing phenomena of stress. [10]

Dr. Meenakshi A. Thalor, Monika S. Budhavant at el., (2019) to propose a framework of machine learning which exploits the characteristics extracted from social network data to accurately identify possible cases of disorder cases. And also proposing a new disorder based tensor model to improve accuracy. To increase the scalability of tensor model, we further improve efficiency with performance guarantees. K nearest neighbours is a simple algorithm that stores all available cases and classifies new cases based on a similarity measure. KNN specially used in statistical estimation and pattern recognition purpose. [11]

B. K. Kiranashree, V. Ambika at al., (2021) this survey exposes stress detection using ML classifiers like Naïve Bayes, Random Forest, and SVM. In this review, to observe the comparison between the machine learning models based on different parameters. Among these three algorithms, SVM gave the best accuracy (96.67%) compared to the naive Bayes and Random forest algorithm but SVM train very less amount of data, NB and Random Forest are not much specific, efficient. [12]

R. Swarna Malika, Ravi at el., (2023) this review can detect the stress by using the machine learning algorithm. Three classification algorithms, logistic regression, random forest, and the decision tree algorithm are applied. The Performance analysis of three algorithms, namely Random forest, Logistic Regression and Decision Tree is performed and accuracy is obtained. When it comes to accuracy, it is observed that the Random forest algorithm outperforms more than the Logistic Regression and Decision Tree algorithms. Hence, by using the Random forest, the prediction of stress detection is performed and the result is obtained [13]

Mochamad Rizky Febriansyah, Nicholas, at el., (2023) this review proposes a machine learning model consisting of Support Vector Machine (SVM), Naïve Bayes, Decision Tree, Random Forest, Bag of Words, and Term Frequency inverse document frequency (TF-IDF) for stress detection. The final evaluation of the model achieved an 80.00% F-1 Score and 75.00% accuracy, and both were scored by SVM. [14]

Chebrolu Naga Harsha Vardhan, DR. V. Ramachandran at el., (2020) this review initially characterize a lot of stress-related printed, visual, and social properties from different angles, at that point convolutional neural networks (CNN) are utilized for subject extraction. Utilizing CNN we can perform feeling analysis of post after the arrangement of the theme. The commitment work is identifying clients who are stressed or not by utilizing the support vector method (SVM). [15]

3. Comparison of Various Techniques Using OSN

REF. No.	TECHNIQUE	POSITIVE IMPACT	NEGATIVE IMPACT	PERFORMANCE
[1]	Machine Learning, Convolutional Neural Network, Natural Language Processing	Both image and text-tweet data are used and the images are calculated with the Optical Character Recognition and the text data are processed by using the Natural Language Processing	When CNN compared to all other classifiers like Naïve Bayes, Maximum Entropy, SVM, MLP it gives improved and more performance in terms of accuracy	By using this approach stress user are predicted earlier and connected to the NGO.

		and Convolutional Neural Network for classifying the tweet content of the user as stressed or non-stressed.		
[2]	Machine Learning, Support Vector Machine, Random Forest)	In this study, we have analyzed how various factors affect the stress of users	Investigate the impact of head movement from outward appearance in the view of complex feelings in data.	It was observed that SVM gives better accuracy (80.2%) than RF.
[3]	Machine learning, TensiStrength, word sense disambiguation.	Using this algorithm accuracy of the prediction is very good.	The major drawback of machine translation is Word Disambiguation.	The proposed method by unigram, bigram and trigram to give better result on ambiguous words.
[4]	Conventional Neural Network, Transductive Support Vector Machine, K-Nearest Neighbours	With the help of these algorithms, we detect the stress level of the person as positive, negative.	Machine learning techniques including the use of various algorithms such as Decision Tree, Naïve Bayes, Random Forest, etc. which gives a lower accuracy of 70% on average	Finally, more accuracy is expected to be achieved using machine learning algorithms and proposed techniques.
[5]	Support vector machines, Naive Bayes algorithm	The proposed system helps predict stress people, thereby helping society in solving serious problems.	It is the use only in the blended learning system phenomena.	Got 74% accuracy with BernoulliNB
[6]	Random Forest (RF), Support Vector Machine (SVM) and Probabilistic Neural Network (PNN)	The Random Forest (RF) model provides higher accuracy rate than Support Vector Machine (SVM) and Probabilistic Neural Network (PNN) models.	The emotion propagation problem in social networks, and found that anger has a stronger correlation among different users than joy, indicating that negative emotions could spread more quickly and broadly in the network.	The RF model outperforms 5-6% higher than other two classification methods.
	Decision Tree,k-mean clustering technique	K-mean clustering techniques in order to cluster and group the	Finding stress level of the user in different states.	In this system, to displayed a system for distinguishing users

[7]		user data and provide accuracy for the user stress levels.		psychological stress states from clients.
[8]	artificial neural network (ANN), deep recurrent neural network (DRNN)	The system's main advantages are its simple design structure, efficient performance, accurate classification, and reduced training model time.	The required attributes are optimally chosen using the alpha-investing strategy to address this problem.	The proposed DRNN technique outperforms the other techniques in terms of accuracy, sensitivity, specificity, precision, recall, and error rate, as evidenced by the results.
[9]	Random Forest and Neural Networks	Revealed that heart rate, heart rate variability, and skin conductance features are the key indicators of stress.	Most studies extracted features before developing AI-based stress detection algorithms	Best AI algorithms for stress detection were found to be Random Forest and Neural Network.
[10]	Factor graph model	In this work, we also discovered several intriguing phenomena of stress.	Preparation starts with simple steps, like loading data, but quickly gets difficult with cleaning tasks that are very specific to the data are working with.	The proposed model can improve the detection performance .With the help of enumeration we build a website for the users to identify their stress rate level and can check other related activities.
[11]	Feature extraction, machine learning	Proposing a new disorder based tensor model to improve accuracy.	It is difficult to identify the disorder because the mental state cannot be observed directly in the registers of online social activities.	To presented a framework for detecting users psychological Mental Disorder states from users monthly social media data, leveraging Face book post content as well as users' social interactions
[12]	Machine Learning, Artificial Neural Network (ANN) classifiers	In this review, mainly focus on different machine learning techniques and physiological parameters for stress detection.	Through this survey it is noticed that they was used a limited and less amount of data. Compared to the above-mentioned techniques Naïve Bayes, Random Forest, and SVM gives the best accuracy but these techniques have not given much specific and efficient result.	As a result, Dynamic BN and support vector machines give the best accuracy of 84.6% and 82.7% respectively.

[13]	Machine learning	Three classification algorithms, logistic regression, random forest, and the decision tree algorithm are applied.	defined Stress is a subjective phenomenon that is difficult to measure comprehensively	By using the Random forest, the prediction of stress detection is performed and the result is obtained
[14]	Support Vector Machine (SVM), Naïve Bayes, Decision Tree, Random Forest	Experiments have achieved good scores, but they are far from perfect, so the development of these personality predictions must be constantly developed to improve their effectiveness.	This study performed clean text in the data pre-processing to clean up the posts in the dataset to ensure the computer could read the texts with some problems or slow timing on decrypting symbols and capitals, making it more efficient and timesaving in the long run.	Final evaluation of the model achieved an 80.00% F-1 Score and 75.00% accuracy, and both were scored by SVM.
[15]	Convolution neural network (CNN)	To fully leverage both content and social interaction information of users' tweets.	Social interaction content, by investigating the content differences between stressed and non-stressed users' social interactions.	proposed model can improve the detection execution

Findings

This analysis indicates that more methods are used to identify the problem of the social media in stress prediction using sentimental analysis. However, an important challenge in detecting stress accurately and reliably is the wide range of stress responses across individuals by using the various techniques. Most of the papers reviewed did not consider these differences among individuals. MACHINE LEARNING (Naive Bayes, Decision Tree, Linear Regression, Artificial Neural Network) are provide solution for that problems. In more accurate solution will move on deep learning based analysis will provide accurate solution of the social media problem and their effects in stress predictions.

Conclusions

In this article, we reviewed 15articles that used wearable technology and sensors to capture physiological data to diagnose stress as one of the important factors that may affect human mental health by using artificial intelligence and machine learning. These papers overall use different pre-processing techniques, different feature extraction techniques, different combinations of measurements and features and machine learning models to enhance the performance of models for better early prediction of stress. The majority of the studies under consideration employed a variety of strategies and approaches to first extract the features before creating models for stress detection utilizing machine learning, Sentimental analysis and various AI algorithms. Furthermore, the best AI algorithms for stress detection were found to be Random Forest and Neural Network. Future work will consider all major issues from social media and provide best solution with deep learning methods.

References

- [1] Preethi Rajam C.R., Uma Maheswari N, "Psychological Stress Prediction on Social Media using Convolutional Neural Network", IJRTE, ISSN: 2277-3878, Volume-8, Issue-2S11, September 2019.
- [2] Akshada kene, Dr. Shubhada thakare, "prediction of mental stress level based on machine learning", 2022
- [3] Reshma Radheshamjee Baheti, Supriya Kinariwala, "Detection and Analysis of Stress using Machine Learning Techniques", IJEAT, ISSN: 2249 – 8958, Volume-9 Issue-1, October 2019.
- [4] S. M. Chaware, Chaitanya Makashir, "Stress Detection Methodology based on Social Media Network: A Proposed Design", IJITEE, ISSN: 2278-3075 (Online), Volume-9 Issue-3, January 2020.
- [5] Mummadi Rama Chandra, Boda Nithin Kumar, "Text based stress detection using machine learning", Volume:05, Issue: 02, February-2023.
- [6] P.Iswarya, D.Maheswari, "Stress detection methods in Social networks: A Comparative study", IRJET, Volume: 07, Issue: 04, Apr 2020.
- [7] S.Venkateswaran, K.Sangeetha, "Human stress detection based on social interactions", IRJET, Volume: 05 Issue: 03, Mar-2018.
- [8] Sameer Dev Sharma, Sonal Sharma, "Deep Recurrent Neural Network Assisted Stress Detection System for Working Professionals", 2022.
- [9] Moein Razavi, Samira Ziyadidegan, "Machine Learning Techniques for Prediction of Stress-Related Mental Disorders: A Scoping Review", 2022.
- [10] Thilagavathi. P, Pushkala. K, "Detecting Stress Based on Social Interactions in Social Network", IJERT, ISSN: 2278-0181, 2018.
- [11] Dr. Meenakshi A. Thalor, Monika S., "Mental Disorder Detection on Social Media using Machine Learning", IJRAT, E-ISSN: 2321-9637, 2019.
- [12] B. K. Kiranashree, V. Ambika, "Analysis on Machine Learning Techniques for Stress Detection among Employees", ISSN: 2249-0701 Vol.10 No.1, pp.35-37, 2021.
- [13] R. Swarna Malika, Ravi, "Stress detection using machine learning techniques", JETIR, ISSN-2349-5162, Volume 10, Issue 3, 2023.
- [14] Mochamad Rizky Febriansyah, Nicholas, "Stress detection system for social media users", 2022.
- [15] Chebrolu Naga Harsha Vardhan, DR. V. Ramachandran, "effective stress detection based on social interactions in social networks", ISSN NO: 0377-9254, Vol 11, Issue 5, May 2020.