International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
ISSN:2147-6799 www.ijisae.org Original Research Paper

A Web Application to Predict Stress via Keyboard Data and Sensor Data

Dr. Meenakshi Thalor[1]*, Dr. Mrunal Pathak[2], Dr. Deeplakshmi Zingade[3],Dr. Vandana Kale[4]

Submitted:14/03/2024 **Revised**: 29/04/2024 **Accepted**: 06/05/2024

Abstract: Stress is a significant problem today, and people may not always be aware of their stress levels. Therefore, it is crucial to identify and recognize stress early and accurately. In literature, various stress detection systems were introduced which results in extensive use of IOT devices. This study proposes a machine learning approach to detect stress, which involves using keyboard data to predict stress levels in computer users. The experiment consisted of two phases, where physiological data was collected and analyzed using a machine learning framework. By analyzing data such as blood pressure and body temperature individuals can avoid stress-related medical conditions. The study assessed the accuracy of stress detection using machine learning algorithms, including Decision Tree Classifier and presents a web application which makes use of keyboard and heart rate sensor.

Keywords: *Decision Tree Classifier, Keyboard, Stress,Sensor*

1. Introduction

Stress is the body's response to a stressful situation that is characterized by feelings of anxiety or pressure. In clinical terms, stress can include severe mental health is-sues such as depression or anxiety attacks. Stress is a psychological and physiological condition that can cause significant discomfort and distress. Stress can result from a mismatch between situational demands and an individual's inability to cope with them. The human body's complex homeostasis is challenged by various states, which trigger distinctive adaptive responses coordinated by the central nervous system, also known as the stress response system. Work pressure, traumatic events, and bereavement are some of the factors that can lead to stress. While everyone experiences some level of stress in their daily lives, it serves as the body's emergency response mechanism. However, excessive stress can become unhealthy and have negative impacts on a person's overall health, emotional well-being, productivity, and quality of life. Prolonged stress can reduce work productivity and result in various physical and mental health problems, such as gastrointestinal, musculoskeletal, cardiovascular, and mental illnesses. Therefore, efforts have been made to predict these diseases before they occur. While past research on stress prediction was con-ducted in a laboratory setting, current research is focused on creating non-invasive strategies using wearable technology. Stress prediction models may not be highly accurate due to the highly subjective and individual nature of stress patterns. Therefore, person-dependent models that are trained using data gathered over a longer timeframe may achieve greater accuracy.

[1,2,3,4] *AISSMS Institute of Information Technology, Pune, Maharashtra, India*
** Corresponding Author Email: meenakshi.thalor@aissmsioit.org*

2. Related Work

According to [1] a brand-new supervised learning model that naturally supports these features for text classification. SS3 was created to be a broad framework for addressing ERD issues. On the CLEF's eRisk2017 pilot challenge for early depression identification, we assessed our model. The majority of the 30 submissions to this competition utilized cutting-edge techniques. For the knowledge base of these intelligent systems, knowledge engineers are typically required to manually code all the facts and regulations obtained from human specialists through interviews (KB). Nevertheless, this manual method is exceedingly costly and prone to mistakes be-cause a genuine expert system's knowledge base (KB) has thousands of rules.

According to [2] the automatic detection of depressive symptoms in text messages from Russian VKontakte users. We outline the process of creating a dataset of user profiles and suggest psycholinguistic and stylistic indicators of depression in literature. We assess machine learning techniques for identifying depressive symptoms in social media posts. accomplished a sadness detection task using text messages from 1020 users of the Russian-language social network Vkontakte. We created a sample of 248 users' posts collections with binary depression/control group classification by examining Beck Depression Inventory scores and processing the raw data. We extracted fresh psycholinguistic features from user writings and formed TF-IDF and dictionary-based feature sets.

According to [3] the scraped data obtained from SNS users is processed using ma-chine learning. Depression may be more easily and effectively detected using Natural Language Processing (NLP), categorized using Support Vector Machine (SVM) and Naive Bayes method. An

isolating hyper plane serves as the formal identifier of a Support Vector Machine (SVM), a discriminative classifier. In other words, the method produces an ideal hyper plane that classifies new models given tagged training data (supervised learning). This hyper plane may be a line that divides a plane into two portions in two-dimensional space, with one class on each side of it. The term "naive Bayes classifiers" refers to a group of characterization methods based on the Bayes theorem. It's a single method, or rather a collection of algorithms, where each algorithm adheres to a common criterion, such as the requirement that any two highlights in a group be independent of one another.

According to [4] a systematic literature review (SLR) is a procedure for locating, evaluating, and interpreting the sources that are accessible in order to provide responses to a number of research questions. It is possible to detect depression early on social media decision tree model is a powerful tool for making predictions and classifications based on complex sets of input data. Due to the existence of specific characteristics in how these subjects use their social media, according to analysis done to address questions about text-based mental illness detection based on the social media activity of people with mental disorders. This SLR discovered that the majority of studies employ deep learning models like RNN on the early diagnosis of depression cases because to the limited availability of data, despite the little quantity of research utilising a text-based technique.

According to [5] Reddit users' postings to see if there are any indicators that might show how relevant online people feel about depression. To do this, we train the data using Natural Language Processing (NLP) methods and machine learning techniques, and then test the effectiveness of our suggested strategy. We find a vocabulary that is more prevalent in narratives of depression. The results demonstrate that the performance accuracy of our suggested strategy may be greatly increased. Bi-gram, along with the Support Vector Machine (SVM) classifier, is the best single feature for detecting depression with 80% accuracy and 0.80 F1 scores. The Multi-layer Perceptron (MLP) classifier has the best performance for depression identification, thereby demonstrating the power and usefulness of the combined features (LIWC+LDA+bigram). In [6][7][8][9][10] the stress is detected using social media text or tweets.

3. Methodology

The system takes input from both real-time data and user input data, including typing speed, error rate, body temperature, and heart rate, which are stored in a CSV file and treated as the dataset. The dataset is pre-processed, split into training and testing data, and used to train a decision tree model.

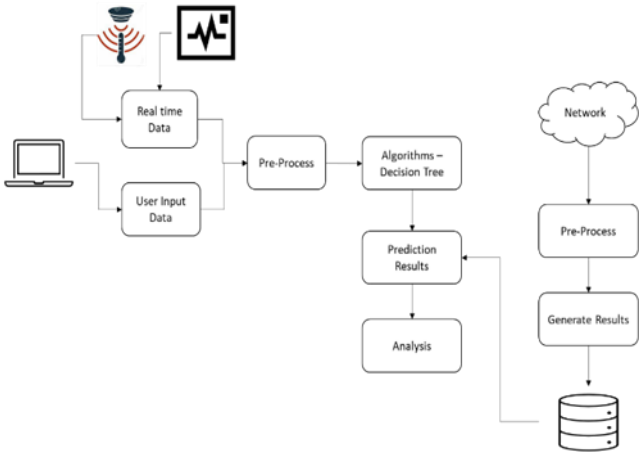

Fig. 1. System Architecture

The performance of the model is evaluated using the testing dataset, and the process is repeated iteratively to develop a robust decision tree model. The trained model is then used to predict stress levels based on user input, and the results are stored in a database and added to the existing CSV file to improve the model's accuracy. The user is provided with feedback based on the predicted stress level.

Decision tree classification algorithm is implemented, as decision tree model is a powerful tool for making predictions and classifications based on complex sets of input data. Assuming the provided dataset, we use the following formula to predict stress using a decision tree algorithm:

stress_level = f (typing_speed, incorrect_words, body_temperature, heart_rate, blood_pressure)

where stress_level is the predicted level of stress, and f is the decision tree algorithm. The decision tree algorithm will use the values of the input parameters to recursively split the data into smaller subsets, based on the most informative features at each step, until it reaches the terminal nodes that correspond to the predicted stress levels.

4. Results and Discussion

This proposed web application makes use of typing speed of user and their health parameter through small sensor which provides temperature and blood pressure data to web application and machine learning algorithm make use of all data to predict stress level. Dataset for stress prediction is prepared by consulting with doctors and refined further as model makes correct predictions. Firstly, the experimental design would involve collecting keyboard data from participants and using it to develop a decision tree algorithm that can predict stress levels. The system is able to give predict severe stress, moderate stress and mild stress/ no stress as shown in table1. The algorithm would then be tested using new data to evaluate its accuracy. The experimental results would likely include metrics such as precision, recall, accuracy, and F1-score to evaluate the

performance of the decision tree algorithm. These metrics would be used to assess how well the algorithm is able to predict stress levels based on keyboard data.

Table 1: Experimental Results

Parameters	Case 1	Case 2	Case 3
Typing speed per min	6.0	34.0	10.0
Incorrect Words	4	7	0
Body Temperature	32.50C	33.00C	32.30C
Heart Rate	97bmp	95bmp	73bmp
Upper BP	120mmHg	150mmHg	110mmHg
Lower BP	72mmhg	95mmHg	80mmHg
Stress Prediction	Severe	Moderate	Mild/no stress

Fig. 2 and fig 3 shows the snapshot of application for stress prediction. Overall, the experimental results provide valuable insights into the accuracy and performance of the algorithm as shown in table 1. These insights could be used to refine the algorithm and improve its effectiveness in predicting stress levels.

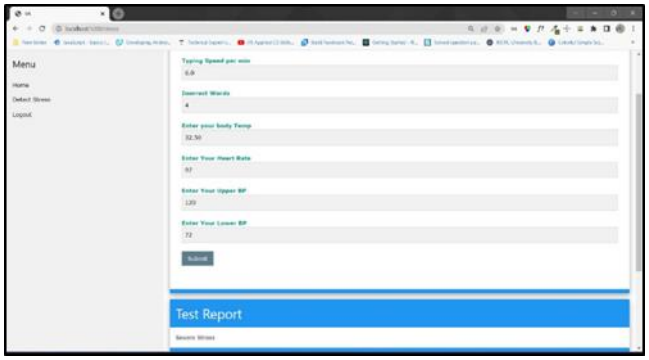

Fig. 2. Parameter inputs for Severe Stress

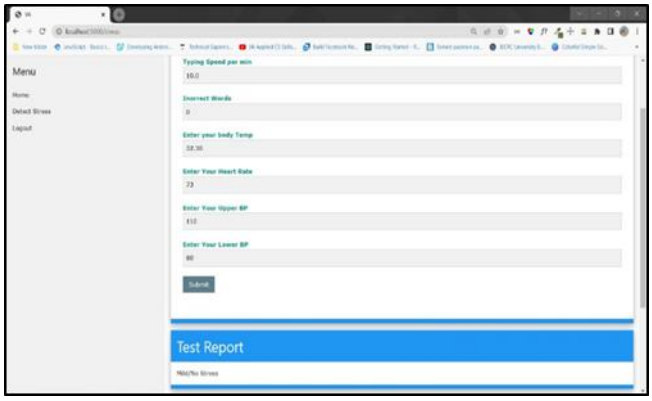

Fig. 3. Parameter inputs for Mild Stress

The experimental results show that the decision tree algorithm has an accuracy of 80%, meaning that it correctly predicts stress levels 80% of the time. The precision of 85%, indicating that when the algorithm predicts stress, it is correct 85% of the time. The recall with 75%, indicating that the algorithm correctly identifies 75% of all instances of stress

Conclusion

This paper proposed a stress prediction mode using decision tree classification which makes use of user typing speed and body parameters like temperature and blood pressure in consideration. In digitization era, most of the time a person spends on typing on keyboard or keypad and using this system a notification about one's stress level can be raise and accordingly user can take preventive action before any serious health issues. This application also recommends the stress relief exercise that one can go for in case of severe stress conditions. Through this approach, controlling blood pressure, body temperature, oxygen levels, and pulse rate, stress indicators and many other such parameters is possible, which are essential to heart health, including vascular age and cardiac index for the same. This system can further enhance by integrating voice notification for stress level.

References

[1] Burdisso, Sergio G., Marcelo Errecalde, and ManuelMontes-y-Gómez. "A text classification framework for simple and effective early depression detection over social media streams." Expert Systems with Applications 133 (2019): 182-197

[2] Stankevich, Maxim, et al. "Depression detection from social media profiles." International Conference on Data Analytics and Management in Data Intensive Domains. Springer, Cham, 2019.

[3] Al Asad, Nafiz, et al. "Depression detection by analyzing social media posts of user." 2019 IEEE International Conference on Signal Processing, Information, Communication & Systems (SPICSCON). IEEE, 2019.

[4] William, David, and Derwin Suhartono. "Text-based depression detection on social media posts: A systematic literature review." Procedia Computer Science 179 (2021): 582-589.

[5] Tadesse, Michael M., et al. "Detection of depression-related posts in reddit social media forum." IEEE Access 7 (2019): 44883-44893.

[6] Shah, Faisal Muhammad, et al. "Early depression detection from social network using deep learning techniques." 2020 IEEE Region 10 Symposium (TENSYMP). IEEE, 2020.

[7] Chiong, Raymond, Gregorious Satia Budhi, and Sandeep Dhakal. "Combining sentiment lexicons and content-based features for depression detection." IEEE Intelligent Systems 36.6 (2021): 99-105.

[8] Narayanrao, Purude Vaishali, and P. Lalitha Surya Kumari. "Analysis of machine learning algorithms for predicting depression." 2020 international conference on computer science, engineering and applications (iccsea). IEEE, 2020.

[9] Laijawala, Vidit, et al. "Classification algorithms based mental health prediction using data mining." 2020 5th International Conference on Communication and Electronics Systems (ICCES). IEEE, 2020.

[10] AlSagri, Hatoon S., and Mourad Ykhlef. "Machine learning-based approach for depression detection in twitter using content and activity features." IEICE Transactions on Information and Systems 103.8 (2020): 1825-1832.