
ISSN: 2321-2152
IJMECE
International Journal of modern
electronics and communication engineering
E-Mail
editor.ijmece@gmail.com
editor@ijmece.com
www.ijmece.com

ISSN:2321-2152 www.ijmece .com

Vol 10, Issue.4 Dec 2022

OPTIMIZATION OF WEARABLE BIOSENSOR DATA FOR STRESS CLASSIFICATION USINGMACHINE LEARNING

[1] Boddupelli Durgabhavani, [2] Tirugulla Neelima,[3] Dr.Doppalapudi Pavan Kumar, [4] Gunashetti Ganesh

[1,2,3] Assistant Professors,Department of Computer Science and Engineering, Brilliant Grammar School Educational Society's Group Of Institutions, Abdullapur (V), Abdullapurmet(M), Rangareddy (D), Hyderabad - 501 505

[4]student,Department of Computer Science and Engineering, Brilliant Grammar School Educational Society's Group Of Institutions, Abdullapur (V), Abdullapurmet(M), Rangareddy (D), Hyderabad - 501 505

ABSTRACT

Examining the efficacy of meditation audio in reducing stress after academic exposure, this research also delves into the usage of wearable sensors for real-time stress monitoring. The MIST records physiological signals, including HRV, BVP, and EDA that are extracted from the IBI. Using Genetic Algorithms and Mutual Information, a hybrid classification strategy is implemented to reduce feature redundancy. The hyperparameters of the machine learning system are then fine-tuned using Bayesian optimisation. Findings show that when EDA, BVP, and HRV are combined, the GB algorithm performs better for 2-level and 3-level stress categorisation. On the other hand, findings from EDA and HRV alone are encouraging as well. Also, according to SHAP Explainable AI (XAI) research, the two most important characteristics for stress categorisation are HRV and EDA. The results provide credence to the idea that meditative music might help alleviate stress. These findings demonstrate the promise of wearable technology integrated with machine learning for the immediate identification and alleviation of academic stress.

I.INTRODUCTION

A growing number of studies have shown that stress, especially in school settings, may have negative effects on children' emotional and physical health. Many medical issues, such as anxiety, depression, and heart disease, may develop as a result of chronic stress. Consequently, enhancing health and productivity necessitates early stress diagnosis and treatment. One potential answer for real-time stress monitoring is the use of wearable biosensors, which can collect physiological data and analyse it to identify stress reactions. Important physiological signals that are indicative of stress are captured by these biosensors. Stress classification methods may not be as accurate when using raw data from wearable devices due to the presence of noise and redundancy. To solve this problem, real-time stress detection may be achieved by processing and classifying the data using machine learning techniques. Furthermore, by determining which elements are most important for categorisation, Explainable AI (XAI) methods may make machine learning models more transparent and the results easier to understand for end users. Using feature selection approaches like Genetic Algorithms and Mutual Information, as well as machine learning techniques like Gradient Boosting, this research aims to optimise data from wearable biosensors for stress categorisation. Additionally, the study investigates the possibility of using Explainable AI, and more especially SHAP analysis, to learn more about the stress-related physiological signals. In order to provide a complete method for both monitoring and controlling stress in academic environments, this research also seeks to understand the effects of meditation audio on stress reduction. This project aims to provide better real-time stress detection and management solutions by combining wearable electronics, machine learning, and XAI.

II.METHODOLOGY

A) System Architecture

Figure 1 shows the process and architecture of the proposed system, which uses machine learning and Explainable AI (XAI) to optimise data from wearable biosensors for stress categorisation in real-time. A centralised server is the brains of this system, coordinating model training, data processing, and aggregation from several wearable devices that record physiological signals including EDA, BVP, and HRV obtained from IBI. The MIST and similar stress-inducing activities record these signals from users so that stress levels may be monitored in real-time.

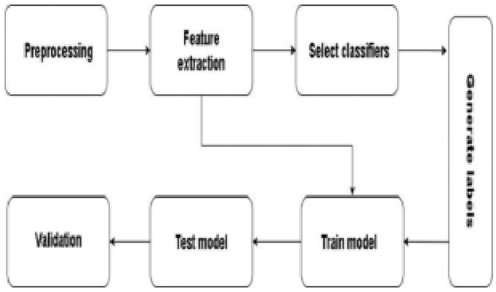

Fig. 1: System Architecture

Clients and servers are the two main parts of the system. Client devices, such biosensors worn by people, gather physiological data and preprocess it to find attributes that machine learning models can exploit. Afterwards, a centralised server receives and processes these characteristics, which include HRV, BVP, and EDA. For stress classification, the server compiles the data and using complex mL techniques including GB, SVM, and LR.

Physiological data is utilised to categorise stress levels using SVM, LR, and GB. While Logistic Regression is used for probabilistic modelling of stress risk, Support Vector Machines are perfect for recognising complicated patterns in high-dimensional data. By merging many underperforming models into one strong one, the ensemble learning technique known as Gradient Boosting is able to increase accuracy. To reliably forecast stress levels in a variety of settings, these machine learning models are trained using data collected from several customers, in this case wearables.

In order to get the most out of the machine learning models, the server employs Bayesian optimisation to tweak their hyperparameters. In addition, the server uses SHAP (Shapley Additive Explanations) and other Explainable AI approaches to make the model's predictions easier to understand. This increases the system's trustworthiness and openness by letting users and healthcare providers know what goes into stress forecasts.

By storing all sensitive user data on the client device, data privacy is preserved throughout the process. In keeping with privacy and confidentiality rules, the server is only informed of processed insights and feature upgrades. To further guarantee the accuracy and robustness of the classification models, the system incorporates techniques to manage possible data preparation issues such unbalanced data, missing values, and outliers.

Returning to the wearable devices, the

trained models are then used to classify stress in real-time. Users may get constant feedback on their stress levels and get suggestions for remedies like meditation or relaxation exercises based on data that's analysed in real-time.

Finally, the suggested method improves stress categorisation using data collected from wearable biosensors by combining machine learning with Explainable AI. This system provides an efficient and scalable solution for stress management and mental health monitoring by integrating real-time physiological monitoring with powerful classification algorithms and guaranteeing interpretability with XAI. Both individual health and healthcare treatments as a whole stand to benefit greatly from the use of machine learning and wearable tech in this setting.

B) Proposed Federated Learning-Based Model

While protecting sensitive information, the Federated Learning (FL) architecture enables clients to train their models locally and communicate changes to the server. The federated architecture allows for the use of many machine learning algorithms for fraud detection, with each algorithm adding to the total model by updating the global model with local insights.

To calculate the local update for Support Vector Machine (SVM), one must first choose the best separating hyperplane, which maximises the margin between valid and fraudulent transactions. Here is one way to depict the change to the FL setting for every client kkk:

∇Lk(Wt) is the gradient of the local loss function at client k, η is the learning rate, and Wt represents the model parameters (support vectors and coefficients). We update the global model by adding together all the local modifications using a weighted total when every client sends their updates to the server.

The goal of training a model for Logistic Regression is to determine the likelihood of a fraudulent transaction. At every client k, the update for the logistic regression parameters θ is:

$$\theta_{k,t+1} = \theta_t - \eta \nabla L_k(\theta_t)$$

The parameter vector θt, the learning rate η, and the gradient of the loss function particular to client k, denoted as ∇Lk(θt), are all defined here. Following the aggregation of updates, the weighted total of all client changes is used to update the global model parameters.

In Gradient Boosting, every client trains its own set of decision trees, with each tree learning from its predecessors and improving upon its mistakes. In this case, the update rule is repeatedly modifying the tree predictions:

$$f_{k,t+1}(x) = f_{k,t}(x) + \eta \cdot h_k(x)$$

In this context, fk,t(x) represents the present forecast for input x at iteration t, hk(x) stands for the recently learnt decision tree model for client k, and η signifies the learning rate. The server then uses the aggregated updates to improve the global model after training.

Using a weighted average of the data points at each client, the server in the FL framework aggregates the local updates for each algorithm. For SVM, Logistic Regression, or Gradient Boosting, the global model Wt+1 is updated as:

$$W_{t+1} = \sum_k \frac{n_k}{n} W_{k,t+1}$$

For all data points, nk is the value. This method guarantees that the global model may use the varied insights from different clients without compromising privacy.

C) Dataset

In order to forecast people's stress levels, this study makes use of a dataset of 10,001 rows of data gathered from wearable biosensors, which capture a variety of demographic and physiological characteristics. User demographics, activity levels, and readings from various sensors are all part of the data set. These factors allow for a more precise categorisation of stress levels by providing a holistic perspective of the physiological states. Machine learning models are trained on this information and then utilise these characteristics to accurately forecast stress and track health issues. This dataset enables multi-dimensional analysis of stress detection by incorporating data points such as heart rate, HRV, skin temperature, electrodermal activity (EDA), activity level, and demographic parameters like age and gender.

HeartRate	HRV	EDA	SkinTemp	ActivityLeve	Stress	Gender	Age
121.378817	31.9783199	1.51062624	36.6903705	low	medium	female	36
69.8146696	42.5277816	0.60893213	36.6445455	medium	high	male	33
73.2534160	74.4576956	0.67160759	36.5246219	high	medium	male	31
75.0080951	54.8292994	0.47801972	36.8562769	low	medium	male	41
80.3716650	51.5364912	0.33229054	36.2682831	medium	medium	male	18
85.7411009	28.8144856	1.09813346	37.0607084	medium	medium	female	18
76.1780727	48.1356358	0.65511470	35.7843337	medium	low	female	18
96.4945608	19.6786606	1.17640648	36.5734997	low	low	male	42
101.832124	10.6289619	1.27066128	36.7519189	medium	high	male	30
74.7222615	43.4067586	0.40098249	36.7259469	low	high	male	32
83.7773784	35.6695959	0.46755674	36.9364867	low	medium	male	52
99.6107803	37.5655122	0.98422980	36.7102693	high	high	female	39
83.0027473	30.8330671	1.55080861	36.8527174	high	medium	female	45
106.282397	33.0865718	1.11851576	37.3365378	medium	medium	male	34
76.8032394	51.0714532	0.55382293	36.7035040	low	medium	male	18

HeartRate is a representation of the user's cardiac activity derived from the heart rate measurements (in beats per minute) recorded by the wearable biosensors.

A person's stress levels, the stability of their autonomic nervous system, and general health may be revealed by measuring their heart rate variability, or HRV.

EDA is a measure of the activity of the sweat glands that rises while we're under stress. Changes in skin conductance, as measured by EDA readings, are associated with levels of emotional arousal.

You may learn about your body's thermal reaction to stress by looking at your skin temperature readings, which might change depending on how much stress you're under.

The user's activity level indicates their current state of physical activity, which might impact their physiological reactions. It measures their activity level and shows whether they are resting, active, or very active.

Based on physiological data, stress is categorised as low, medium, or high, and it serves as the goal variable in the dataset that indicates the stress level.

Because people's physiological reactions to stress might vary depending on their gender, the demographic characteristic "gender" can be useful for this kind of analysis.

The user's age is a significant component in stress prediction since people's reactions to and ability to cope with stress might alter as they get older.

D) Feature Selection

By simplifying the data and zeroing down on the most important characteristics, feature selection is a crucial step in making machine learning models work as well as possible. Selecting the most important characteristics guarantees accurate predictions in this study, which uses data from wearable biosensors to assess stress levels. Statistical tests such as chi-square and ANOVA are used to evaluate the relevance of characteristics after correlation analysis has been used to find duplicate features. To further assess the significance of features, we use embedded approaches via decision tree-based algorithms and wrapper methods like Recursive Feature Elimination (RFE). The correlation between characteristics and stress levels may be better understood with the use of mutual information. The dataset is fine-tuned for enhanced model performance after choosing essential characteristics including HeartRate, HRV, EDA, and ActivityLevel. This results in less overfitting and increased accuracy.

III.CONCLUSION

Using machine learning and explainable AI (XAI), this study shows how to optimise data from wearable biosensors for stress categorisation. Incorporating sophisticated feature selection methods allows us to train the model with just the most relevant data, which improves accuracy while simplifying the process. Combining XAI with machine learning methods like Gradient Boosting not only improves stress classification performance, but also makes the results easier to grasp and analyse. The findings highlight the promise of wearable tech and AI for managing stress in real-time, which is particularly useful in professional and academic contexts. Further, this study demonstrates how explainable AI may improve personalised stress detection and management systems by making decision-making more transparent and trustworthy.

IV.REFERENCES

1. P. S. Prabu, "A study on academic stress among higher secondary students", *Int. J. Humanities social Sci. Invention*, vol. 4, no. 10, pp. 63-68, 2015.
2. S. D. Ghatol, Academic Stress Among School Students, Chennai, India:Allied Publishers, 2019.
3. A. Waqas, S. Khan, W. Sharif, U. Khalid and A. Ali, "Association of academic stress with sleeping difficulties in medical students of a Pakistani medical school: A cross sectional survey", *PeerJ*, vol. 3, pp. e840, Mar. 2015.
4. V. M. Bhujade, "Depression anxiety and academic stress among college students: A brief review", *Indian J. Health Wellbeing*, vol. 8, no. 7, pp. 1-26, 2017.
5. S. Maji, A. Chaturmohta, D. Deevela, S. Sinha, S. Tarsolia and A. Barsaiya, "Mental health consequences of academic stress amotivation and coaching experience: A study of India's top engineering undergraduates", *Psychol. Schools*, vol. 61, no. 9, pp. 3540-3566, Sep. 2024.
6. S. Byun, A. Y. Kim, E. H. Jang, S. Kim, K. W. Choi, H. Y. Yu, et al., "Detection of major depressive disorder from linear and nonlinear heart rate variability features during mental task protocol", *Comput. Biol. Med.*, vol. 112, Sep. 2019.
7. M. Jafari, A. Shoeibi, M. Khodatars, S. Bagherzadeh, A. Shalbaf, D. L. García, et al., "Emotion recognition in EEG signals using deep learning methods: A review", *Comput. Biol. Med.*, vol. 165, Oct. 2023.
8. L. Holtz, M. Martinez, K. Paton, K. Rosich and E. Schnittka, "Effects of physiological stress response on short-term memory recall", *J. Adv. Student Sci.*, 2017.

9. C. Anders and B. Arnrich, "Wearable electroencephalography and multi-modal mental state classification: A systematic literature review", *Comput. Biol. Med.*, vol. 150, Nov. 2022.

10. A. Arsalan and M. Majid, "Human stress classification during public speaking using physiological signals", *Comput. Biol. Med.*, vol. 133, Jun. 2021.

11. M. Ullah, S. Akbar, A. Raza and Q. Zou, "DeepAVP-TPPred: Identification of antiviral peptides using transformed image-based localized descriptors and binary tree growth algorithm", *Bioinformatics*, vol. 40, no. 5, pp. 305, May 2024.

12. S. Akbar, Q. Zou, A. Raza and F. K. Alarfaj, "IAFPs-Mv-BiTCN: Predicting antifungal peptides using self-attention transformer embedding and transform evolutionary based multi-view features with bidirectional temporal convolutional networks", *Artif. Intell. Med.*, vol. 151, May 2024.

13. S. Akbar, A. Raza and Q. Zou, "Deepstacked-AVPs: Predicting antiviral peptides using tri-segment evolutionary profile and word embedding based multi-perspective features with deep stacking model", *BMC Bioinf.*, vol. 25, no. 1, pp. 102, Mar. 2024.

14. A. Raza, J. Uddin, A. Almuhaimeed, S. Akbar, Q. Zou and A. Ahmad, "AIPs-SnTCN: Predicting anti-inflammatory peptides using fastText and transformer encoder-based hybrid word embedding with self-normalized temporal convolutional networks", *J. Chem. Inf. Model.*, vol. 63, no. 21, pp. 6537-6554, Nov. 2023.

15. S. Akbar, A. Raza, T. A. Shloul, A. Ahmad, A. Saeed, Y. Y. Ghadi, et al., "PAtbP-EnC: Identifying anti-tubercular peptides using multi-feature representation and genetic algorithm-based deep ensemble model", *IEEE Access*, vol. 11, pp. 137099-137114, 2023.