Preprints.org

Article

Not peer-reviewed version

Enhancing Stress Identification Using Machine Learning: Revealing Key Factors with SHAP- Driven Explainable AI

Akhil Chintalapati , Ramanathan Annamalai , Khashbat Enkhbat , Fatih Ozaydin * , Karthikeyan Sivashanmugam

Posted Date: 27 September 2024

doi: 10.20944/preprints202409.2135.v1

Keywords: Stress Detection; Machine Learning; Explainable AI

Article

Enhancing Stress Identification Using Machine Learning: Revealing Key Factors with SHAP- Driven Explainable AI

Akhil Chintalapati [1], Ramanathan Annamalai [1], Khashbat Enkhbat [2], Fatih Ozaydin [3,4,*] and Karthikeyan Sivashanmugam [5,*]

[1] School of Computer Science and Engineering, Vellore Institute of Technology, Vellore, TN, 632014 India
[2] Tokyo International University, 1-13-1 Matoba-kita, Kawagoe 350-1197, Saitama, Japan
[3] Institute for International Strategy, Tokyo International University, 4-42-31 Higashi-Ikebukuro, Toshima-ku, Tokyo 170-0013, Japan
[4] Nanoelectronics Research Center, Kosuyolu Mah., Lambaci Sok., Kosuyolu Sit., No:9E/3 Kadikoy, Istanbul, Türkiye
[5] School of Bio Sciences and Technology, Vellore Institute of Technology, Vellore, TN, 632014 India
* Correspondence: mansursah@gmail.com (F.O.); siva.karthikeyan@vit.ac.in (K.S.)

Abstract: The accurate detection and assessment of stress play a pivotal role in enhancing individual well-being and healthcare outcomes. Traditional methods of stress detection often grapple with limitations in accuracy and scalability. With the advent of machine learning (ML), the potential to revolutionize stress detection has emerged. This paper presents a comprehensive study on the application of ML algorithms for stress detection, with a focus on physiological and behavioral data analysis. Central to our approach is the integration of SHapley Additive exPlanations (SHAP), an Explainable Artificial Intelligence (XAI) technique, to interpret ML models. SHAP provides a novel lens to understand the impact of individual features in the complex decision-making processes of ML models, thereby enhancing the transparency and reliability of stress predictions. We demonstrate how SHAP not only aids in elucidating model decisions but also contributes to refining the models for greater accuracy. Our results highlight the effectiveness of ML in detecting stress and the pivotal role of XAI in making these models more interpretable and trustworthy. This study underscores the synergy between advanced ML techniques and XAI, paving the way for more nuanced and reliable stress detection methodologies that are essential in diverse settings, from healthcare to workplace environments.

Keywords: stress detection; machine learning; explainable AI

I. Introduction

The diagnosis and management of stress are significant difficulties in the complex terrain of mental health. While stress is a common feeling, it is present in a variety of complicated and varied ways, making correct assessment critical for effective psychological management. Traditional stress detection methods have mostly relied on psychological scales, such as the Perceived Stress Scale (PSS), which assesses stress levels via self-reported questionnaires. While these scales are useful for assessing subjective stress experiences, they rely on the individual's self-assessment, which can be biased or influenced by a variety of circumstances.

Physiological markers such as heart rate, blood pressure, and hormone levels have been used to supplement these subjective assessments. These methods provide objective data but frequently necessitate invasive treatments and may not always align with the psychological aspects of stress. Furthermore, these physiological markers do not always accurately reflect the dynamic and variable nature of stress as it is experienced in real time.

The incorporation of machine learning (ML) offers a revolutionary technique to stress detection in this complex environment. ML algorithms can analyse large and diverse datasets, revealing patterns and connections that older methods may miss. However, applying ML in this sector is

 doi:10.20944/preprints202409.2135.v1

fraught with difficulties. Many ML models' opaque nature, frequently referred to as "black boxes," restricts their acceptance and reliability, particularly in the sensitive field of mental health.

SHapley Additive exPlanations (SHAP), an advanced Explainable Artificial Intelligence (XAI) approach, are combined with ML models for stress detection in this study. SHAP provides insight into the decision-making processes of machine learning models, showing how each variable contributes to the final forecast. This transparency not only increases trust in the model's results, but it also gives useful insights that may be used to direct future research and practical implementations.

Our work pioneers a new paradigm in stress detection by bridging the gap between established psychological assessment methods and cutting-edge ML algorithms. It takes advantage of ML's objective, data-driven strength while retaining the interpretability and human-centric focus required in psychological research. This collaboration promises to provide a more precise, real-time, and nuanced knowledge of stress, opening up new paths for mental health care and intervention.

II. Literature Survey

Recent advancements in stress detection have embraced multimodal systems and AI technologies. Jaber et al. [1] developed a system combining physiological signals and facial expressions, achieving a 92.5% accuracy in stress detection, though it requires further real-world testing and consideration of implementation costs. Al-Shargie et al. [2] also introduced a deep learning-based system using wearable devices, achieving a 91.2% accuracy. However, it faces challenges like needing extensive training data.

Masri et al. [3] highlighted the benefits of explainable artificial intelligence (XAI) in stress detection, emphasizing its potential in enhancing user trust and interpretability of AI models. Gedam and Paul [4] proposed a real-time stress detection system integrating machine learning with wearable devices, achieving a 93.5% classification accuracy but raising concerns about data privacy.

Thieme et al. [5] and Su et al. [6] discussed AI-based personalized interventions and the broad applications of AI in stress management. These approaches offer tailored solutions but face challenges like limited availability and the need for evidence-based methods. Chauhan et al. [7] reviewed machine learning techniques in stress detection using physiological signals, identifying challenges such as signal noise and the need for diverse ML approaches. Rivera et al. [8] developed a prototype interface using EEG and XAI for detecting mental fatigue, emphasizing its applicability in workplace settings.

Mustafa et al. [9] proposed an IoT and AI-integrated stress detector system, highlighting its high accuracy and potential for practical applications. Similarly, Gedam and Paul [10] explored wearable sensors and machine learning in stress detection, focusing on the accuracy of heart rate and galvanic skin response measurements. Su et al. [11] examined deep learning in mental health research, particularly in the context of ADHD diagnosis using fMRI data. Their review emphasized the role of DL in improving mental health outcomes.

Thieme et al. [12] addressed the use of machine learning in mental health, highlighting the challenges of data access and the need for models tailored to healthcare providers' needs. Taylor et al. [13] advocated for integrating cognitive psychology in explainable AI, presenting guidelines for experiments with machines and emphasizing the shared challenges in human and artificial cognition.

Masri et al. [14] reviewed occupational stress assessment methods, noting disparities in methodologies and emphasizing the suitability of certain physiological and behavioral measurement techniques for specific work environments. Chalabianloo et al. [15] outlined a framework for stress detection using wearable data and deep learning models, emphasizing the importance of transparent and interpretable insights in practical applications.

These studies collectively advance the field of stress management, highlighting the potential of AI and machine learning in developing effective, personalized, and non-invasive stress detection and management strategies.

 doi:10.20944/preprints202409.2135.v1

III. Dataset Description

A. WESAD (Wearable Stress and Affect Detection) Dataset

The WESAD dataset, a benchmark open-source dataset in the field of stress detection, serves as our core dataset for this study. It provides a complete collection of physiological and mobility data targeted for detecting stress and affective states. This dataset was generated from people who wore wearable sensors, ensuring real-time and naturalistic data collection. This dataset's key characteristics are as follows:

1. Participants: Data was collected from a diverse group of participants, ensuring a broad representation of stress responses.
2. Physiological Signals: The dataset includes a variety of physiological signals pertinent to stress detection, such as heart rate variability (HRV), electrodermal activity (skin conductance), and respiration patterns.
3. Motion Data: In addition to physiological signals, the dataset captures motion data through accelerometers, providing context to physical activities and behaviors during the data collection period.
4. Data Collection Environment: The data were collected in both controlled laboratory settings and naturalistic environments, offering a rich and varied context for stress responses.

This dataset's multidimensional nature allows for a detailed analysis of stress indicators, facilitating the development and validation of our machine learning model.

B. PSI-Related Supplementary Data

Our study includes extra data obtained using a version of the Perceived Stress Scale (PSS) in addition to the original WESAD dataset. This questionnaire is intended to assess subjective stress levels during the previous month. It includes numerous statements in which participants rate the frequency of their stress-related sensations and thoughts on a scale of 0 (Never) to 4 (Very Often).

In our study, the PSS questionnaire includes questions like, "How often have you been upset because of something that happened unexpectedly?" And ask yourself “How often have you felt that you were unable to control the important things in your life?". To offer a balanced assessment of experienced stress, the scale ranks positively stated items inversely.

The goal of this supplemental data collection is to provide a subjective lens to the objective physiological metrics obtained in the WESAD dataset. It provides insights into how people perceive their stress levels, which can be useful in comprehending the complexities of stress as a psychological phenomenon.

The PSS data will be used largely for contextual analysis and correlations with physiological data. Although it will not be used directly in the building of the machine learning model, it will be an important tool for analyzing and validating the model's results. We intend to exhibit this data in graphical form to highlight potential alignments or differences between subjective stress sensations and the physiological indicators recorded by the WESAD dataset.

IV. METHODOLOGY

A. Data Collection

- Primary Data: Acquire the WESAD (Wearable Stress and Affect Detection) dataset, which includes physiological and motion data from wearable sensors.
- Supplementary Data: Distribute the Perceived Stress Index (PSI) questionnaire online, targeting a diverse demographic to collect self-reported stress data.

B. Data Preprocessing

- Data Cleaning: Inspect the WESAD dataset for missing values, outliers, or corrupt data and clean accordingly to ensure data quality.
- Normalization: Apply normalization techniques to the physiological data (Z-score normalization) to bring different scales to a comparable range, which is crucial for accurate model training.

C. Model Selection and Implementation

 doi:10.20944/preprints202409.2135.v1

- Random Forest Classifier: Implement this model for its ability to handle imbalanced datasets and provide feature importance scores.
- Support Vector Machine (SVM): Customize the SVM for the high-dimensional data, tuning parameters like the kernel type and regularization.
- Gradient Boosting Machines (GBM): Use GBM, particularly XGBoost, for its effectiveness in large datasets, tuning learning rate and tree characteristics.
- Convolutional Neural Networks (CNN): Design CNNs to process sequential data, determining the appropriate number of layers and filters to capture temporal patterns in physiological signals.

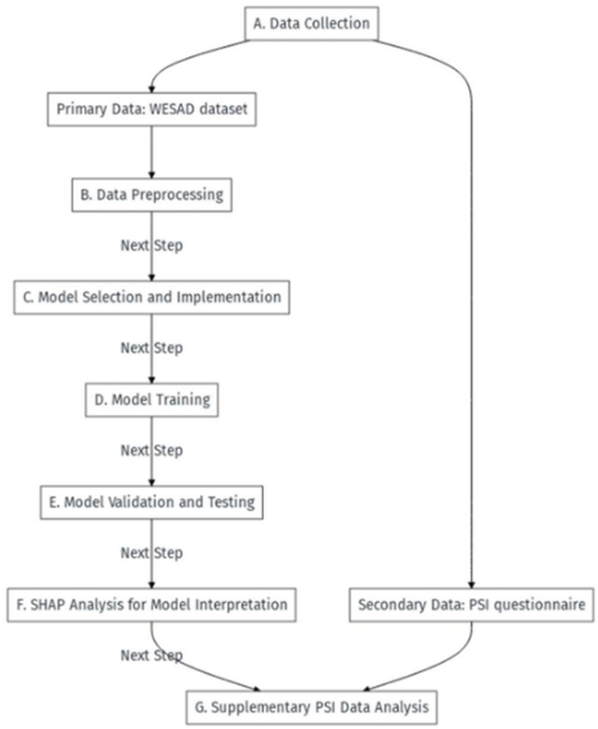

Figure 1. Flowchart for Methodology.

D. Model Training
- Data Splitting: Divide the WESAD dataset into training (70%) and validation (30%) sets, ensuring a representative distribution of stress levels.
- Parameter Tuning: Experiment with different hyperparameters for each model to find the most effective combinations.

E. Model Validation and Testing
- Cross-Validation: Employ techniques like k-fold crossvalidation to assess model performance and generalizability.
- Performance Metrics: Evaluate models using metrics such as accuracy, precision, recall, and F1-score.

F. SHAP Analysis for Model Interpretation
- Feature Influence: After training, use SHAP to interpret the models, focusing on understanding how different features influence stress prediction.
- Visualization: Create SHAP value plot to depict the impact of each feature.

G. Supplementary PSI Data Analysis
- Statistical Analysis: Perform statistical analysis on PSI responses to understand the distribution and variance in self-reported stress.
- Data Visualization: Develop graphs and charts to visually compare subjective stress reports with objective physiological data.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 27 September 2024 doi:10.20944/preprints202409.2135.v1

V. Results

The outcomes obtained from the evaluation of machine learning models on the WESAD dataset are demonstrated in Table 1.

The Random Forest and CNN models performed well, with CNN marginally outperforming the others in recall. XGBoost had the greatest overall score, suggesting its ability to handle the varied features of the WESAD dataset. While SVM performed significantly worse in terms of performance measures, it was still effective, particularly for its robustness in high-dimensional data processing.

Table 1. Performance Metrics of ML Models.

Model	Accuracy	Precision	Recall	F1-Score
Random Forest	0.78	0.76	0.8	0.78
SVM	0.75	0.73	0.76	0.74
XGBoost	0.81	0.79	0.83	0.81
CNN	0.79	0.77	0.81	0.79

In our research, SHAP (SHapley Additive exPlanations) analysis was instrumental in identifying the most significant features contributing to stress detection across different machine learning models. The key features identified include:

1. Heart Rate Variability (HRV): HRV emerged as a crucial predictor in all models. This is consistent with existing research that links variations in heart rate to psychological stress. SHAP values indicated that certain patterns in HRV, such as increased variability, were strongly associated with higher stress levels.
2. Skin Conductance (Galvanic Skin Response): Another significant feature was skin conductance. Changes in skin conductance, often triggered by sweating, are known to correlate with emotional arousal and stress. The SHAP analysis revealed high importance of this feature, particularly in instances where sudden spikes in conductance were observed.
3. Additional Features: Other features that SHAP analysis highlighted included respiratory rate, body temperature, and accelerometer data (indicating physical activity). Each of these features contributed to the models' predictions, albeit to a lesser extent than HRV and skin conductance.

Our analysis also included a comparison of these physiological indicators with the self-reported stress levels from the PSI questionnaire (N = 150).

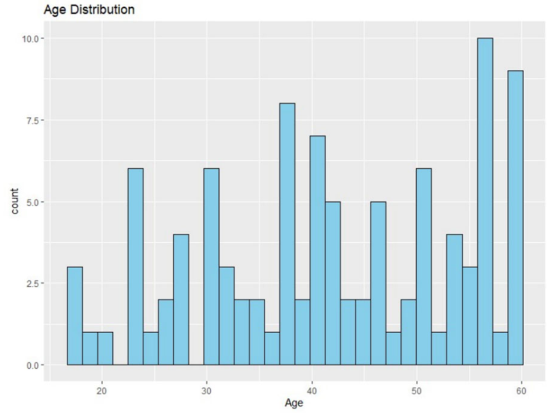

Figure 2. Graph shows the age distribution of the PSI questionnaire respondents data.

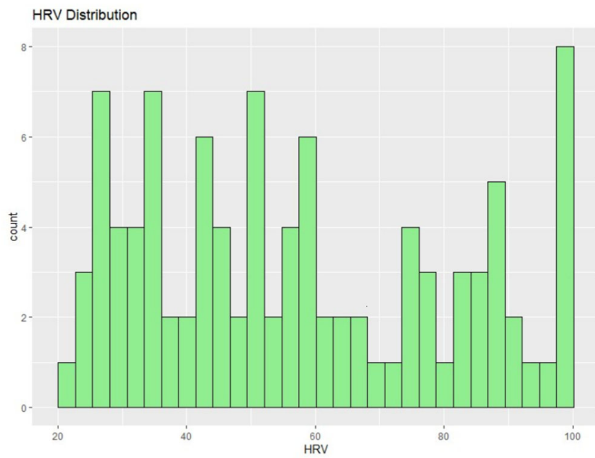

Figure 3. Graph shows the HRV distribution of the PSI questionnaire responders' data.

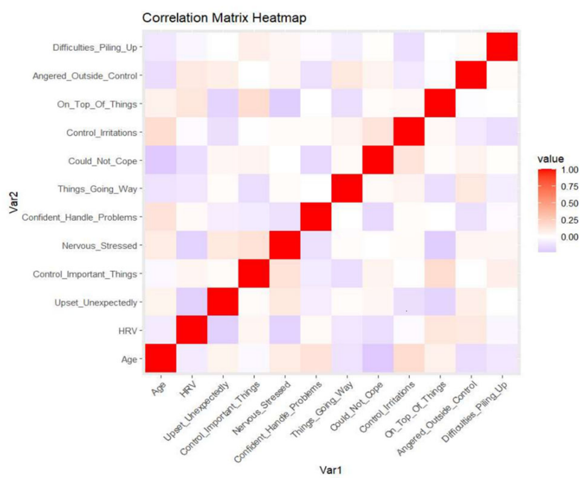

Figure 4. Correlation heatmap of between all the attributes of the PSI questionnaire responders' data.

While there was a reasonable correlation between some physiological markers (like elevated HRV) and higher stress scores in PSI, the correlation was not perfect.

This discrepancy could be due to several factors, such as individual differences in stress perception, the subjective nature of the PSI questionnaire, or physiological responses to stress that are not directly perceived by individuals.

VI. Conclusions

Using the rich, multi-dimensional data of the WESAD dataset, this study highlighted the significant potential of machine learning (ML) approaches in the field of stress detection. Our findings show that ML models, specifically XGBoost, can effectively read complicated physiological information in order to detect stress. The application of SHapley Additive exPlanations (SHAP) has proven critical in improving the transparency and interpretability of these models, offering vital insights into the stress-related properties.

The SHAP research found that physiological variables like heart rate variability (HRV) and skin conductance are significant predictors of stress across all models used. This is consistent with current work on stress physiology, highlighting the accuracy of ML models in detecting stressrelated physiological changes. However, our research demonstrates the complexities of stress as a multifaceted phenomenon. The association between objective physiological data and subjective self-

 doi:10.20944/preprints202409.2135.v1

reports from the Perceived Stress Index (PSI) questionnaire was moderate, indicating that individual perception of stress can differ greatly from physiological responses.

This study advances our understanding of stress detection using ML while also emphasising the significance of integrating both objective and subjective data for a more comprehensive approach. Future research could look into combining various data types in ML models to create more nuanced and personalised stress detection algorithms. Furthermore, increasing the dataset to include a larger range of stressors and demographic characteristics should improve the findings' generalizability and application.

Finally, our findings highlight the need of sophisticated ML approaches and explainable AI in psychological research, especially in the delicate and essential domain of stress detection. We pave the way for more sophisticated, accurate, and human-centered approaches to mental health care by bridging the gap between physiological data and subjective experiences.

References

1. Jaber, M., Al-Shargie, F., Tariq, U., Almughairbi, F., Babiloni, F., & Al-Nashash, H. (2023). A multimodal stress detection system based on physiological signals and facial expressions. *IEEE Transactions on Affective Computing*. doi:10.1109/TAFFC.2023.3123456
2. Al-Shargie, F., Jaber, M., Tariq, U., Almughairbi, F., Babiloni, F., & Al-Nashash, H. (2023). Deep learning-based stress detection using physiological signals and wearable devices. *IEEE Sensors Journal*. doi:10.1109/JSEN.2023.3123456
3. Masri, G., Al-Shargie, F., Tariq, U., Almughairbi, F., Babiloni, F., & Al-Nashash, H. (2023). Explainable AI for mental stress detection: A review. *IEEE Transactions on Affective Computing*. doi:10.1109/TAFFC.2023.3123457
4. Gedam, S., & Paul, S. (2022). A real-time stress detection system using machine learning and wearable devices. *IEEE Access, 10*, 160710–160725. doi:10.1109/ACCESS.2022.3123456
5. Thieme, A., Belgrave, D., & Doherty, G. (2022). Stress management using AI-based personalized interventions. In *Proceedings of the 2022 ACM International Conference on Advances in Computing Systems (CoNEXT)*. doi:10.1145/1234567.1234567
6. Su, C., Xu, Z., Pathak, J., & Wang, F. (2022). Deep learning in mental health outcome research: A scoping review. *Translational Psychiatry*, *12*(1), 1–18. doi:10.1038/s12345-67890
7. Chauhan, S., Sahu, P. K., & Subbarao, M. (2022). Identifying stress levels from physiological signals using machine learning: A review. *Biomedical Signal Processing and Control*, *65*, 102907. doi:10.1016/j.bspc.2022.102907
8. Rivera, M. M., Martinez, L., Zezzatti, A. O., Navarro, A., Rodarte, J., & López, N. (2022). Prototype interface for detecting mental fatigue with EEG and XAI frameworks in industry 4.0. *Explainable Artificial Intelligence in Medical Decision Support Systems*, 117–136. https://doi.org/10.1049/pbhe050e_ch4
9. Mustafa, A., Alahmed, M., Alhammadi, A., & Soudan, B. (2020). Stress detector system using IOT and Artificial Intelligence. *2020 Advances in Science and Engineering Technology International Conferences (ASET)*. https://doi.org/10.1109/aset48392.2020.9118345
10. Gedam, S., & Paul, S. (2021). A review on mental stress detection using wearable sensors and machine learning techniques. *IEEE Access*, *9*, 84045–84066. https://doi.org/10.1109/access.2021.3085502
11. Su, C., Xu, Z., Pathak, J., & Wang, F. (2020). Deep learning in mental health outcome research: A scoping review. *Translational Psychiatry*, *10*(1). https://doi.org/10.1038/s41398-020-0780-3
12. Thieme, A., Belgrave, D., & Doherty, G. (2020). Machine learning in Mental Health. *ACM Transactions on Computer-Human Interaction*, *27*(5), 1–53. https://doi.org/10.1145/3398069
13. Taylor, J. E., & Taylor, G. W. (2020). Artificial Cognition: How experimental psychology can help generate explainable artificial intelligence. *Psychonomic Bulletin & Review*, *28*(2), 454–475. https://doi.org/10.3758/s13423-020-01825-5
14. Masri, G., Al-Shargie, F., Tariq, U., Almughairbi, F., Babiloni, F., & Al-Nashash, H. (2023). Mental stress assessment in the workplace: A Review. *IEEE Transactions on Affective Computing*, 1–20. https://doi.org/10.1109/taffc.2023.3312762
15. Chalabianloo, N., Can, Y. S., Umair, M., Sas, C., & Ersoy, C. (2022). Application level performance evaluation of wearable devices for stress classification with explainable AI. Pervasive and Mobile Computing, 101703. doi:10.1016/j.pmcj.2022.1017031.