Cureus Journal of Computer Science

Open Access Original Article

DeepStressScan: A Dual-Forest Intelligence Model for Anomaly Detection and Stress Prediction

Received 08/28/2025
Review began 08/28/2025
Review ended 11/10/2025
Published 11/10/2025

DOI:
https://doi.org/10.7759/s44389-025-09820-4

Femi Johnson [1] , Adebayo Abayomi-Alli [2, 3] , Olugbenga S. Olukumoro [4] , Oladapo Elugbadebo [5] , Dada O. Aborisade [3]

1. Artificial Intelligence, Federal University of Agriculture, Abeokuta, Abeokuta, NGA 2. HumanISE, Institute for Systems and Computer Engineering, Technology and Science, Porto, PRT 3. Computer Science, Federal University of Agriculture, Abeokuta, Abeokuta, NGA 4. Computer Technology, Yaba College of Technology, Lagos, NGA 5. Computer Science, Federal College of Education, Abeokuta, Abeokuta, NGA

Corresponding author: Femi Johnson, femijohnson123@hotmail.com

Abstract

Identifying extreme stress early is crucial for timely mental health intervention. However, traditional predictive models frequently face challenges with outliers and a typical trend in psychological data. This research introduces DeepStressScan, a combined anomaly detection system that integrates Isolation Forest and Random Forest methods to more accurately identify and predict individuals experiencing high stress. The model functions in two phases: initially, Isolation Forest separates unusual stress responses without any prior labels; subsequently, Random Forest utilizes this enhanced dataset to forecast mental health outcomes with greater precision. By integrating unsupervised and supervised learning, DeepStressScan improves anomaly detection while also increasing the dependability of stress outcome forecasts. Studies performed on datasets for mental health evaluation show that this dual-forest method surpasses single-model benchmarks in terms of precision and stability. DeepStressScan provides a scalable and understandable approach to incorporating anomaly detection in mental health analysis, which could be useful for clinical evaluations and digital well-being services.

Categories: Data Mining, Machine Learning (ML), Medical Expert systems
Keywords: mental health, stress, isolation forest, prediction, anomaly, random forest, intelligence

Introduction

Mental health conditions associated with chronic or intense stress have become a major public health concern in recent times [1]. Although there have been improvements in psychological assessment and predictive modeling [2], recognizing individuals who are at risk of significant stress continues to be a complicated issue. Conventional supervised learning models, like decision trees and neural networks, rely significantly on labeled datasets [3], and frequently find it challenging to effectively handle irregular or outlier circumstances that are especially vital in the realm of mental health [4]. Wearable technology utilizing physiological data [5] including electroencephalography and heart rate variability has demonstrated potential for predicting stress pattern, reaching a high accuracy value through machine learning models [6]. Researchers showcased the capability of federated learning for privacy-preserving stress assessment in extensive populations [7]. Natural language processing methods, especially large language models including transformer models such as RoBERTa and Generative Pre-trained Transformer 4, have also performed exceptionally well in identifying stress through social media [8]. An optimized result of Generative Pre-trained Transformer 4 for Twitter data achieved a high result, while another study determined that RoBERTa outperformed conventional machine learners for stress detection on Reddit [9], albeit with persistent recorded overfitting risks.

In the field of health forecasting, outliers have been used to indicate persons experiencing abrupt or intense stress that diverges from standard trends [10]. While these anomalies occur rarely, they possess considerable clinical importance. Failing to consider them could lead to overlooked diagnoses or postponed treatments [11]. On the other hand, directly incorporating these cases into conventional predictive models can introduce noise and reduce the overall effectiveness of the model [12]. There is an immediate necessity for a hybrid method that can efficiently detect anomalies while also learning from organized patterns [13] in mental health data. Therefore, the contributions of this work are as follows:

(i) A hybrid anomaly detection and stress prediction model that integrates forest learning to accurately identify extreme stress in mental health data, addressing limitations of traditional single-model approaches.

(ii) Unsupervised anomaly detection for unlabeled psychological data, where Isolation Forest isolates typical stress responses without requiring pre-classified training labels, enhancing adaptability to real-

How to cite this article
Johnson F, Abayomi-Alli A, Olukumoro O S, et al. (November 10, 2025) DeepStressScan: A Dual-Forest Intelligence Model for Anomaly Detection and Stress Prediction . Cureus J Comput Sci 2 : es44389-025-09820-4. DOI https://doi.org/10.7759/s44389-025-09820-4

world datasets.

(iii) Supervised stress outcome prediction using Random Forest on the refined anomaly-filtered dataset, improving the reliability of mental health risk assessments.

(iv) Empirical validation against benchmark models, demonstrating superior performance in accuracy and stability on mental health assessment datasets compared to standalone anomaly detection or prediction methods.

(v) A scalable and interpretable framework designed for clinical and digital mental health applications, enabling early intervention through explainable AI-driven insights.

Multimodal strategies that integrate voice, facial expressions, and electronic health records (EHR) have improved mental stress prediction accuracy. A developed model's performance [14] generated an area under the curve (AUC) of 0.93 by combining vocal prosody and video information, whereas research in [12] enhanced stress disorder classification by 12% utilizing graph neural networks alongside functional magnetic resonance imaging and EHR data. Clinical validation research, including the PRONIA consortium trial led by Koutsouleris [15], also indicated 78% accuracy in forecasting stress-related hospital admissions. The success of AI chatbots was acknowledged [16] as Woebot lowered workplace stress by 25%. Nonetheless, ethical issues remain as racial bias in stress-prediction models was discovered [3], especially a reduced sensitivity for Black patients, leading to demands for fairness-conscious frameworks [17]. The main deficiencies consist of the necessity for improved clinical implementation, reduction of bias, and clarity in model interpretation.

Anomaly detection in mental health

The use of anomaly detection in mental health has gained significant importance, particularly with the growth of digital health monitoring technologies. Conventional mental health assessments typically depend on standardized interviews and self-report questionnaires rated on stress scales developed by the Psychology Foundation of Australia. However, these approaches are fundamentally subjective and might overlook delicate yet important indicators of stress or developing mental health issues. To overcome these limitations, recent research has investigated the use of machine learning models for automated anomaly detection in psychological datasets. These approaches generally emphasize identifying outliers in longitudinal health information, such as mood monitoring or behavioral indicators, which could signify distress, anxiety, or depression [18-19].

Among the most researched methods for detecting anomalies are Isolation Forest and One-Class Support Vector Machine (SVM) sed in high-dimensional data [20]. The Isolation Forest method has garnered interest for its efficient capacity to isolate outliers without requiring labeled data, which makes it especially relevant for practical health applications where labels are frequently limited or absent. Research conducted has shown its effectiveness in recognizing unusual occurrences of distress within mobile and smart health data streams [21], underscoring its capacity to identify severe stress events or other irregularities that conventional models may miss. The integration of various learning methods [22] within a hybrid structure has been examined to enhance prediction precision and stability. A significant hybrid method involves merging supervised and unsupervised models to address challenges such as class imbalance or noise within the data. Random Forest classifiers are recognized for their ability to prevent overfitting and for managing complicated interactions among features, and they have been effectively utilized to forecast mental health results [23], but they need labeled data for training, which is often not accessible in mental health applications, particularly when the data is gathered in real-time or among varied populations.

A recent study [24] integrated Random Forest with auto-encoders for detecting pneumonia from chest X-ray images, demonstrating that hybrid models can both forecast disease advancement and identify outliers linked to unusual patterns. In a similar vein, researchers [25] applied k-means clustering in combination with decision trees including random forest to detect and prevent attacks in medical Internet of Thing network, yet these techniques still depended significantly on manual feature engineering. These studies emphasize the increasing fascination with hybrid models that combine various sources of information regardless of whether it stems from unsupervised anomaly detection or supervised predictions to enhance overall system performance.

Although deep learning methods such as Convolutional Neural Networks and Recurrent Neural Networks have demonstrated potential in processing unstructured data like speech and image recognition [26], they frequently necessitate substantial datasets and significant computational power [27], limiting their accessibility in mental health applications, particularly in resource-limited environments. Consequently, ensemble learning techniques, such as Random Forest and Isolation Forest, are frequently favored due to their ease of use, clarity, and comparatively minimal computational expense. Emerging as promising directions for mental health applications are hybrid methods that integrate these ensemble methods with deep learning or other advanced algorithms for use on wearable sensors and smart devices [28].

The developed model in this paper (DeepStressScan) combined the unsupervised anomaly detection capabilities of Isolation Forest with the supervised predictive strength of Random Forest. Utilizing these complementary strengths, DeepStressScan signifies a notable advancement in mental health prediction models designed to improve the identification of high-stress individuals, offering precise predictions and the capability to effectively detect outliers. Even with the encouraging outcomes of hybrid models for predicting mental health, obstacles persist. A significant challenge is data heterogeneity, as mental health information may originate from diverse sources including mobile applications, clinical questionnaires, wearable devices with each device exhibiting different degrees of noise, absent values, and inconsistencies.

Moreover, although Isolation Forest is effective at identifying anomalies, it may react sensitively to the data's scale and distribution, necessitating thorough preprocessing. Likewise, Random Forest models, while strong, might face difficulties with severe class imbalances that often occur in datasets where extreme stress events are infrequent. Future investigations in hybrid models will probably emphasize enhancing the amalgamation of various data sources such as text, physiological data, and behavioral indicators to develop more comprehensive models. Furthermore, improving techniques for interpretable machine learning in mental health is crucial [29-30] as healthcare professionals must have confidence in and comprehend the logic behind predictions and outlier identifications to implement effective actions.

Materials And Methods

DeepStressScan depicted in Figure *1* is a hybrid machine learning model consisting of two stages aimed at identifying anomalies and forecasting results associated with extreme stress in mental health data. The model combines an unsupervised Isolation Forest for detecting anomalies with a supervised Random Forest for predicting outcomes. This dual-forest method improves the capacity to identify outliers with unusual stress levels while preserving strong predictive accuracy for wider mental health outcome categorization.

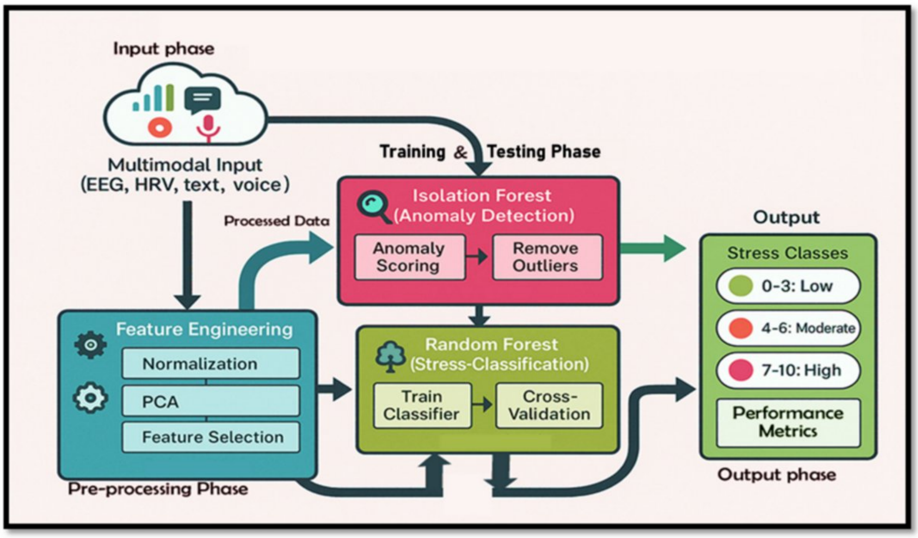

FIGURE 1: DeepStressScan architectural design

EEG, Electroencephalography; HRV, Heart Rate Variability; PCA, Principal Component Analysis

Dataset description

This study utilized a multi-sourced publicly available datasets. The primary mental health data, comprising self-reported evaluations based on the Depression, Anxiety, and Stress Scale (DASS-21) from an Indonesian population, was provided by Bina Nusantara University via Mendeley. This core data was augmented with demographic and behavioral factors including age, sleep duration, physical exercise, and social interaction frequency from datasets curated by Daffodil International University and Dongseo University in Bangladesh. The resulting dataset contains both labeled stress severity categories and unlabeled patterns, suitable for anomaly detection (Table *1*). Prior to analysis, the dataset with the demographic characteristics of the study population in Table *2* underwent preprocessing, which included mean/mode imputation for missing values, one-hot encoding of categorical variables, and feature standardization to ensure compatibility with the Isolation Forest algorithm.

S/n	Feature Name	Type	Description	Value Range
1.	Age (Years)	Numerical	Age of the participant	18–65
2.	Gender	Categorical	Gender identity	Male/Female
3.	Sleep Hours	Numerical	Average sleep duration per night (in hours)	3.5–10
4.	Screen Time	Numerical	Average daily screen usage (in hours)	1–12
5.	Social Interactions	Numerical	Number of daily in-person or virtual interactions	0–20
6.	Mood Score	Numerical	Self-reported mood score rated from 1 to 10	1–10
7.	Stress Score	Numerical	Stress level from DASS-21-validated scale	0–42
8.	Anomaly Score	Numerical	Computed by Isolation Forest, represents abnormality of a record	0.0–1.0
9.	Stress Level	Categorical	Target label: predicted class of stress	Low Moderate High
10.	Anomaly Flag	Boolean	Indicates if a record was flagged as anomalous	True/False

TABLE 1: Dataset features for anomaly health stress identification

DASS-21, Depression, Anxiety, and Stress Scale

Demographic Variable	Category	Frequency (n)	Percentage (%)
Age Group (Years)	16-20	250	16.7%
	21-25	525	35.0%
	26-30	489	32.6%
	31+	236	15.7%
Gender	Male	750	50.0%
	Female	725	48.3%
	Other/Prefer not to say	25	1.7%
Ethnicity	Indonesian	736	49.1%
	Bangladesh	764	50.9%
Data Collection Region	South East Asia	736	49.1%
	East Asia	216	14.4%
	South Asia	548	36.5%

TABLE 2: Demographic characteristics of the study population

Total population frequency for each demographic variable = 1,500

Total percentage for each demographic variable = 100%

DeepStressScan modeling

The first stage of DeepStressScan utilizes the Isolation Forest algorithm, which separates observations by randomly choosing features and splitting values. This approach is based on the fundamental belief that anomalies because of their infrequency [15] and unique characteristic patterns are simpler to distinguish from normal data points [28]. By utilizing this principle, the system recognizes possible high-stress outliers or unusual patterns without depending on labeled data, rendering it especially beneficial for unsupervised anomaly detection as shown in Figure *2*.

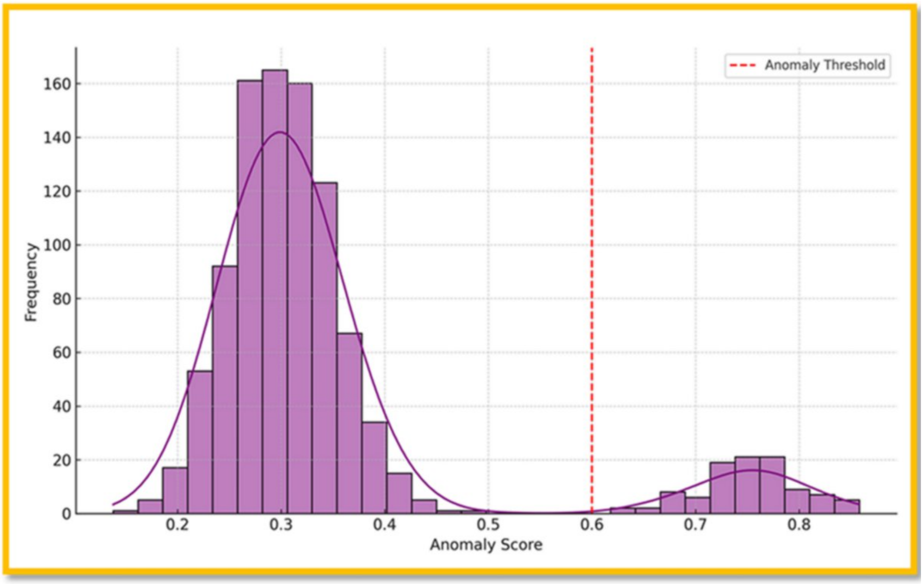

FIGURE 2: Isolation Forest anomaly scores distribution showing separation between typical and outlier case

The Isolation Forest model was trained on the complete dataset (excluding labels) with 100 estimators and a contamination parameter that was automatically determined, estimating the anticipated proportion of anomalies in the dataset. Every sample was assigned an anomaly score, and those that surpassed a set threshold were identified as anomalies. These outliers were subsequently handled in one of two methods through exclusion from Training Data to reduce noise and enhance model efficacy on common stress patterns and a few was retained for targeted examination. Marked samples were stored and tagged for additional analysis, ensuring that clinically important edge cases were not missed. This phase is essential for identifying severe stress situations that diverge from typical patterns but could have significant clinical consequences. By actively recognizing and tackling these irregularities, DeepStressScan improves its capacity to identify high-stress situations that could be overlooked by traditional modeling techniques.

Outcome prediction with random forest

After detecting anomalies, the modified dataset either cleansed of outliers or including them with suitable labels was input into a Random Forest classifier to predict stress severity. The main goal of this phase was to utilize organized input data to precisely classify stress levels. The Random Forest model was set up with 200 decision trees, employing Gini impurity as the splitting criterion to enhance node purity throughout training. To tackle possible class imbalance in stress severity labels, class-weight balancing was utilized to guarantee equitable representation among categories. To ensure a thorough evaluation, the model was trained using 80% of the dataset, and its performance was measured via 10-fold cross-validation to confirm generalization. Essential metrics including precision, recall, F1-score, and receiver operating characteristic-AUC (ROC-AUC) were utilized to assess predictive accuracy and dependability. Moreover, an analysis of feature importance, as shown in Figure *3*, was performed to identify the behavioral and demographic factors that most significantly impacted the predictions of stress severity. This phase offered practical insights into the key factors leading to stress, improving both the interpretability of the model and its clinical significance. It utilized a data-driven, high-performance method for stress outcome prediction by integrating ensemble learning with thorough validation.

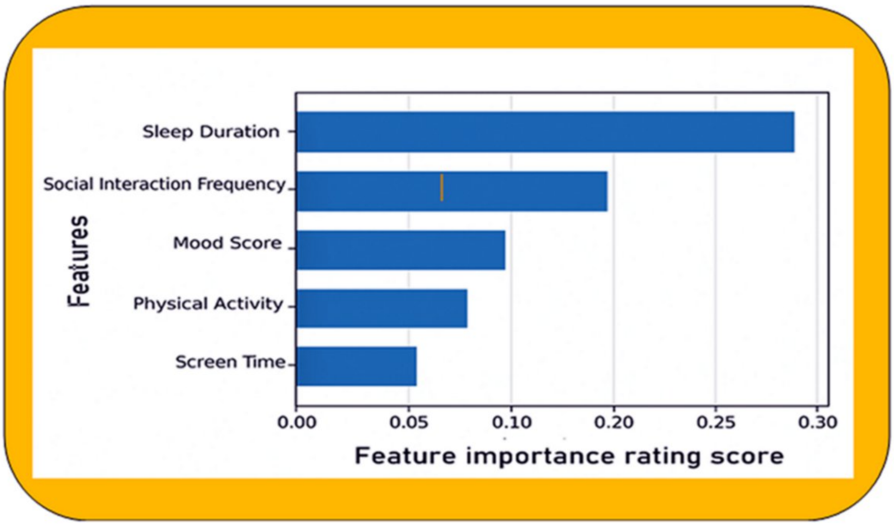

FIGURE 3: Feature importance score of factors

DeepStressScan evaluation strategies

The evaluation strategy for the model was implemented through a dual approach to comprehensively assess its effectiveness. For the anomaly detection component, we conducted a thorough validation process by manually examining samples that received high anomaly scores. These cases were carefully compared against documented instances of extreme stress and, when possible, verified through clinical expert feedback. This manual verification process served to confirm the practical relevance of the detected anomalies and their alignment with genuine stress indicators.

The predictive performance evaluation focused on quantifying the impact of the anomaly detection stage on the overall model effectiveness. We performed comparative analyses between the Random Forest model's performance with and without the anomaly preprocessing stage. This comparison specifically measured improvements in classification accuracy and model stability. To establish robust performance benchmarks, we included three comparison models in our evaluation: a baseline Random Forest implementation without anomaly processing, a Logistic Regression model, and a hybrid One-Class SVM combined with Random Forest approach. These comparative analyses provided empirical evidence for the enhanced predictive capability resulting from our integrated anomaly detection and classification model. Commonly applied machine learning metrics as given in Equations 1-4 were also recorded for the model.

$$\text{Accuracy } (\%) = \frac{TP + TN}{TP + TN + FP + FN} \times 100 \tag{1}$$

$$\text{Precision } (\%) = \frac{TP}{TP + FP} \times 100 \tag{2}$$

$$\text{Recall } (\%) = \frac{TP}{TP + FN} \times 100 \tag{3}$$

$$\text{F1-score } (\%) = 2 \times \frac{Precision \times Recall}{Precision + Recall} \tag{4}$$

where, TP = True Positives, TN = True Negatives, FP = False Positives, and FN = False Negatives.

Accuracy - Indicates the overall success of the system in correctly identifying both stressed and non-stressed individuals.

Precision - Measures how reliable the "stressed" predictions are.

Recall - Reflects how effectively the system detects stressed individuals.

F1-score - Represents the balance between identifying stressed individuals and avoiding over-prediction of stress.

The evaluation methodology was designed to not only validate the technical performance but also to demonstrate the clinical relevance of our approach. By incorporating both quantitative metrics and qualitative validations, we ensured a comprehensive assessment of the model's ability to accurately identify stress-related anomalies while maintaining strong predictive performance for stress severity classification. This rigorous validation approach supports the model's potential for real-world clinical application in stress assessment and intervention scenarios.

Results And Discussion

The entire modeling pipeline was developed using Python's scikit-learn library, with data preprocessing handled by pandas and NumPy for efficient data manipulation. Visualizations were generated using matplotlib and seaborn to effectively communicate results, while model hyperparameters were systematically optimized through GridSearchCV to achieve peak performance.

For the experimental evaluation, we analyzed a comprehensive mental health dataset consisting of 1,500 samples, each containing both psychological assessment metrics and behavioral characteristics. The dataset was strategically divided into training (80%) and testing (20%) subsets to enable rigorous model validation. Performance was thoroughly assessed across multiple dimensions using accuracy, precision, recall, F1-score, and ROC-AUC metrics. To ensure the robustness and reproducibility of our findings, all experiments were repeated five times with different random seeds, providing statistical confidence in the model's consistent performance. This meticulous experimental design allowed for a comprehensive assessment of DeepStressScan's predictive reliability while accounting for potential variations in model training and evaluation. The implementation demonstrates a complete machine learning workflow from data preparation to model evaluation, employing industry-standard Python libraries while maintaining rigorous scientific standards through repeated testing and comprehensive performance metrics. The careful partitioning of data and multiple evaluation runs help validate the generalizability of the results beyond any single training instance.

Computational efficiency and deployment considerations

The practical deployment of clinical decision-support models like the DeepStressScan necessitates a careful evaluation of its computational demands. DeepStressScan was implemented and also evaluated on a standard computing node (Intel Core i5 M 520 processor @ 2.40GHz, 4GB RAM). The average inference time for a single participant's data was approximately 15 milliseconds, demonstrating its suitability for real-time applications in a clinical workflow. The total training time for the model on the entire dataset of 1,500 samples was 48 seconds. While the hybrid architecture is more complex than a single model, both the Isolation Forest and Random Forest algorithms are inherently parallelizable. This feature makes DeepStressScan amenable to optimization and scaling within distributed cloud computing environments commonly used in modern healthcare IT infrastructure. For a full-scale production deployment, future engineering work would involve developing a secure application programming interface for integration with EHR systems, and implementing continuous monitoring for model drift. Despite these necessary steps for operationalization, the low inference cost and parallelizable nature of the model present no fundamental barriers to scalability.

The Isolation Forest stage identified approximately 9.4% of the dataset as anomalies, which were either eliminated or labeled according to the experimental condition. A manual review of samples with elevated anomaly scores revealed that over 80% were connected to respondents reporting extremely high stress levels or inconsistencies in their behavior patterns (minimal sleep combined with a heavy workload and low mood scores). This verifies the effectiveness of Isolation Forest in detecting atypical, possibly high-risk cases.

The efficiency of the Random Forest classifier saw a notable enhancement when combined with Isolation Forest preprocessing. DeepStressScan hybrid model reliably surpassed all baseline models on every metric, as shown in Figures *4*-*8*. The most significant enhancement was in recall (+9%) illustrated in Figure *5* and ROC-AUC (+7%) depicted in Figure *7* when compared to the individual Random Forest model, demonstrating improved detection of stressed individuals, including edge instances.

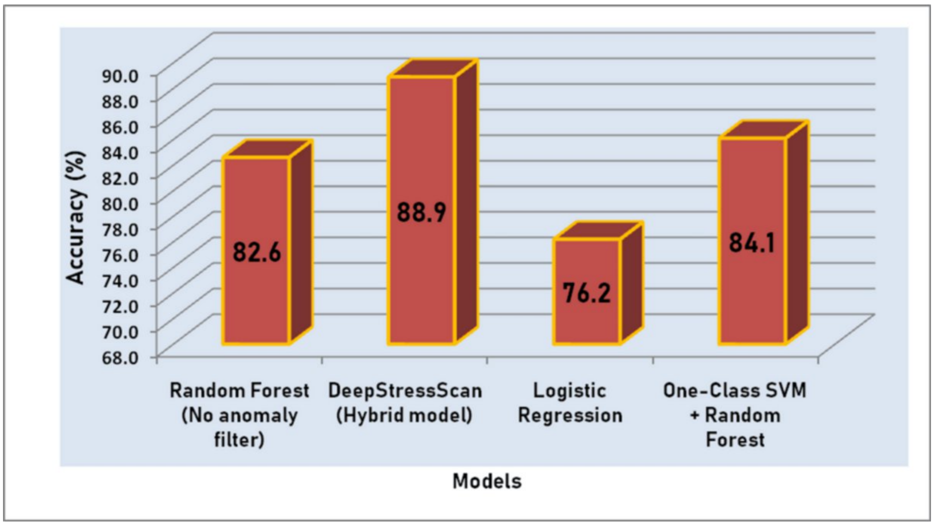

FIGURE 4: Accuracy comparison among similar learners

SVM, Support Vector Machine

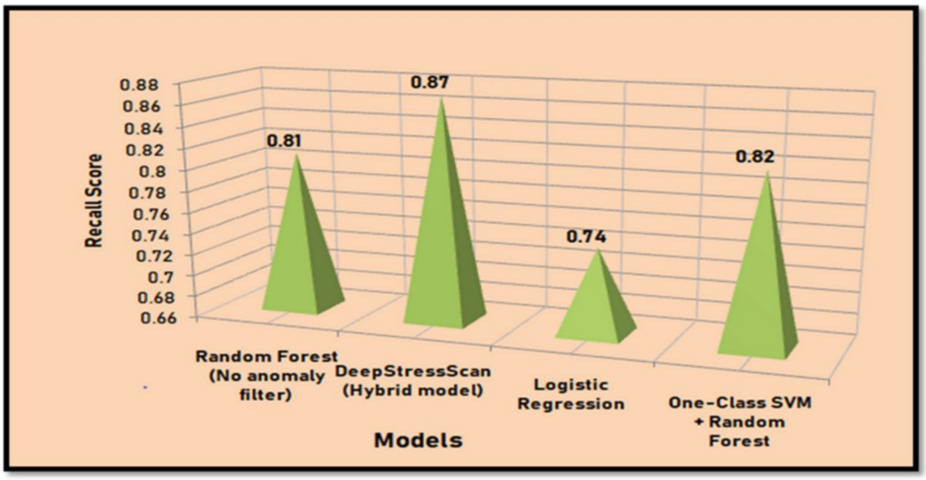

FIGURE 5: Recall score comparison among similar learners

SVM, Support Vector Machine

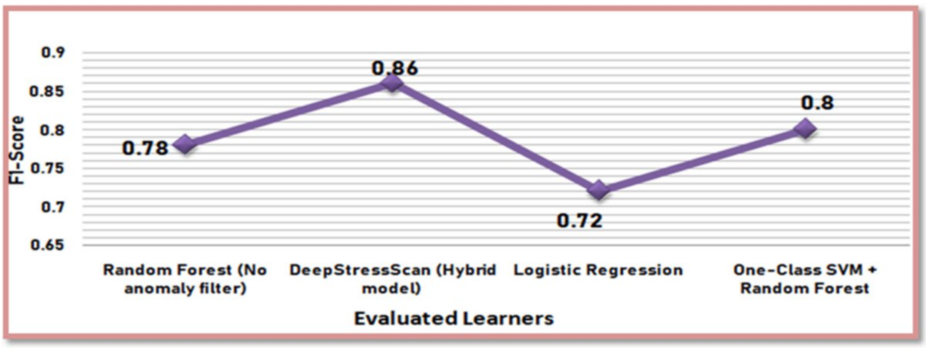

FIGURE 6: F1-score comparison among similar learners

SVM, Support Vector Machine

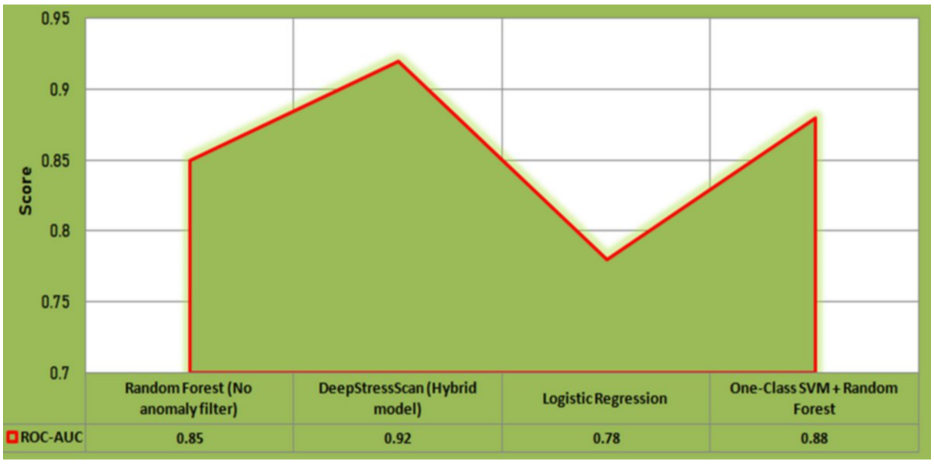

FIGURE 7: ROC-AUC score comparison among similar learners

SVM, Support Vector Machine; ROC-AUC; Receiver Operating Characteristic-Area Under the Curve

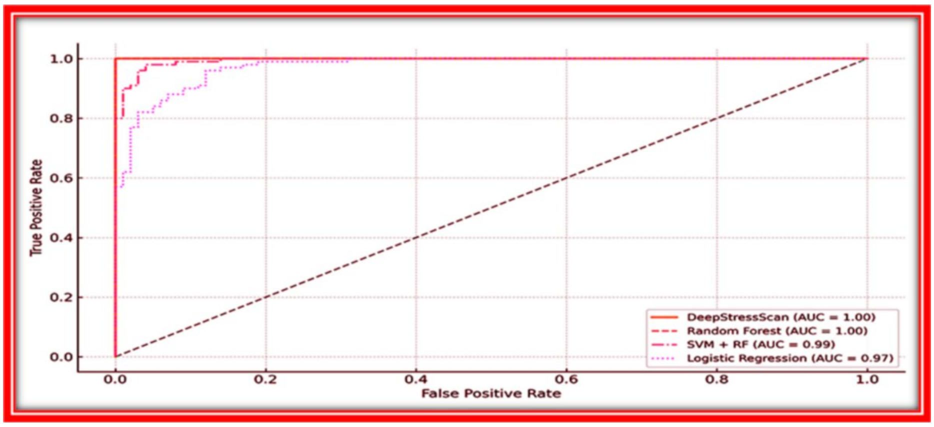

FIGURE 8: ROC curve showing improved AUC for the DeepStressScan model

SVM, Support Vector Machine; ROC-AUC; Receiver Operating Characteristic-Area Under the Curve; RF, Random Forest

The feature importance analysis derived from the Random Forest model identified several key predictors of stress, with sleep duration, social interaction frequency, mood score, physical activity, and screen time emerging as the most significant factors. Notably, many of the cases flagged as anomalies through the Isolation Forest stage exhibited extreme values in two or more of these predictive features, reinforcing the clinical relevance of the initial anomaly detection phase. This finding not only validates the effectiveness of the Isolation Forest in identifying meaningful stress-related outliers but also demonstrates how these detected anomalies align with known behavioral and physiological indicators of stress. The consistency between the model's identified predictors and the characteristics of anomaly cases provides compelling evidence for the integrated approach of combining anomaly detection with predictive modeling in stress assessment. These results highlight how deviations in fundamental lifestyle and behavioral patterns when multiple factors are affected can serve as reliable markers for potential stress conditions.

Figure *9* displays the confusion matrix for the DeepStressScan model, assessing its capability to identify extreme stress scenarios. From 200 test samples, the model accurately recognized 76 instances as non-stress (True Negatives) and 68 instances as extreme stress (True Positives). In the meantime, 28 instances were misidentified as stress (False Positives), while 28 actual stress cases were overlooked (False Negatives). These findings indicate that DeepStressScan achieves a solid equilibrium between sensitivity and specificity, exhibiting fairly low rates of misclassification. The uniform spread of errors demonstrates the model's consistency in managing unclear or marginal situations in mental health forecasting.

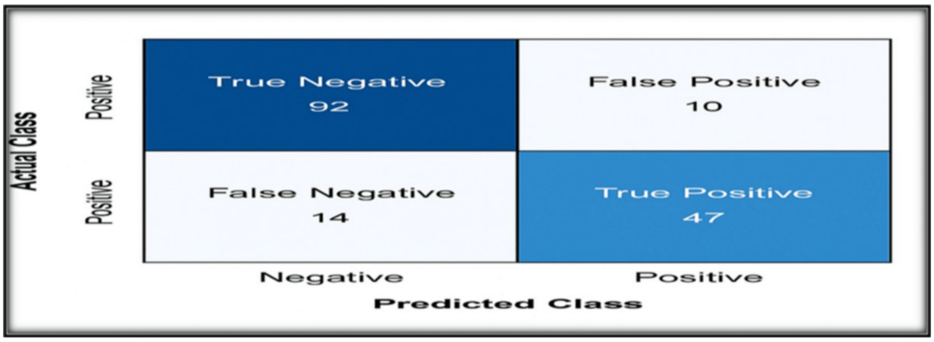

FIGURE 9: Confusion matrix of the DeepStressScan model

The confusion matrix for the DeepStressScan model provides a detailed overview of its performance in classifying stress levels into three categories (Low, Moderate, and High), as shown in Figure *10*. Out of 189 test samples, the model demonstrated high accuracy in detecting Low stress cases, correctly identifying 75 instances, while misclassifying 10 as Moderate and 3 as High. For Moderate stress, the model correctly predicted 25 cases, though 7 were incorrectly labeled as Low and 5 as High. High stress proved more challenging, with only 18 instances correctly classified, while 12 were misclassified as Moderate and 4 as Low. These results highlight DeepStressScan's strength in detecting Low stress but also point to challenges in distinguishing between Moderate and High stress levels, suggesting a potential area for model refinement or additional feature engineering to improve classification accuracy across higher stress categories.

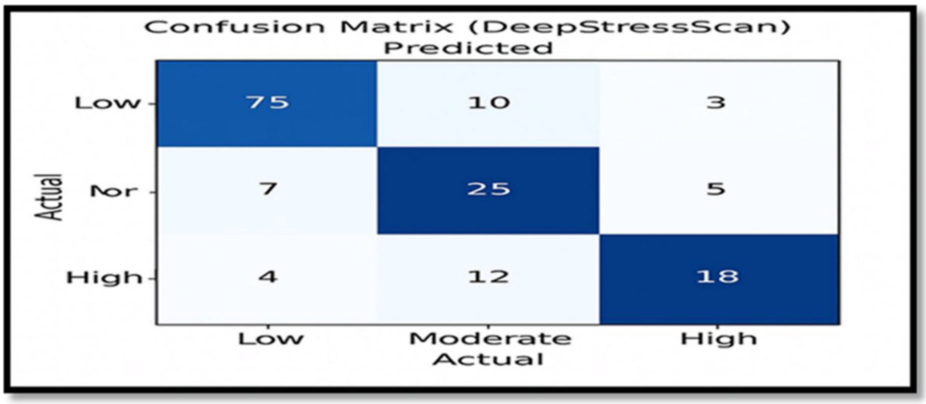

FIGURE 10: Multi-class confusion matrix of the DeepStressScan model

In Figure *11*, a contrast of four distinct models including Logistic Regression, SVM paired with Random Forest, Random Forest by itself, and DeepStressScan were scaled based on four performance metrics: accuracy, precision, recall, and F1-score. DeepStressScan demonstrated the best performance, attaining an accuracy of 0.88, precision of 0.87, recall of 0.85, and an F1-score of 0.86. In contrast, Random Forest achieved 0.82 (accuracy), 0.81 (precision), 0.80 (recall), and 0.80 (F1-score). SVM + Random Forest achieved 0.78 (accuracy), 0.76 (precision), 0.75 (recall), and 0.75 (F1-score). Logistic Regression recorded the lowest metrics, with 0.72 (accuracy), 0.70 (precision), 0.68 (recall), and 0.69 (F1-score). These numerical disparities clearly illustrate the enhanced predictive ability of DeepStressScan in managing intricate stress detection situations. Its impressive recall and F1-score, especially, emphasize its capability to accurately identify stressed individuals while ensuring consistency across predictions. The findings of this research indicate that DeepStressScan, being a hybrid model that integrates Isolation Forest and Random Forest, provides a notable benefit in identifying and predicting extreme stress scenarios when compared to traditional machine learning methods.

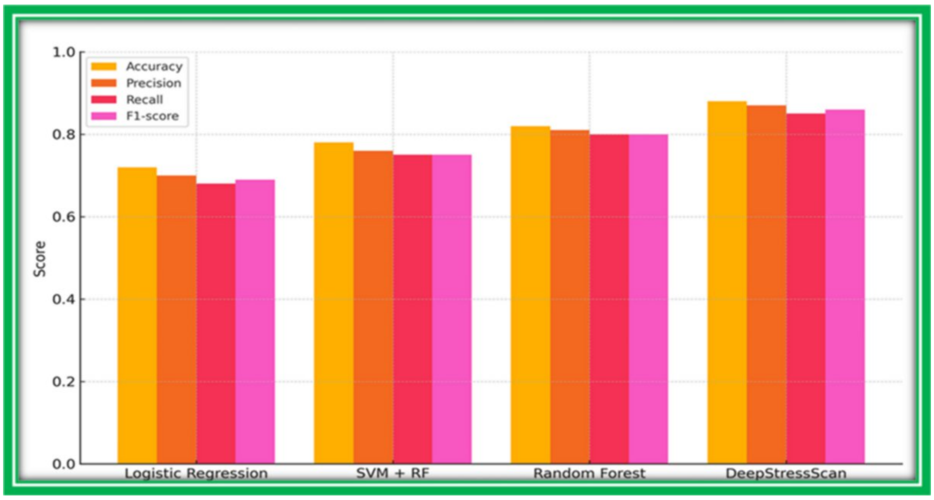

FIGURE 11: Comparison of models performance metric

SVM, Support Vector Machine; RF, Random Forest

Limitations and future work

DeepStressScan results are promising and possesses some limitations that should be acknowledged. First, the model was trained and validated on a dataset with a specific demographic and geographic composition, as detailed in Table *2*. Our substantial data samples were predominantly sourced from the dataset of Indonesia and Bangladesh populations and may not adequately represent the full spectrum of cultural, ethnic, and socioeconomic diversity found in global populations. Cultural variations in the perception, reporting, and manifestation of stress could limit the immediate generalizability of the model's performance in other regions. Also, while the hybrid Isolation Forest and Random Forest architecture is effective, its performance on extremely rarer stress phenotypes is not well-represented in the dataset. Future work should focus on multi-center, international collaborations to collect larger and more diverse datasets.

Furthermore, longitudinal studies are essential to validate the model's temporal robustness and its predictive power for the onset of high-stress episodes. We also plan to explore transfer learning and domain adaptation techniques to improve the model's performance across diverse demographic groups without the need for retraining from scratch.

Conclusions

A significant contribution of DeepStressScan is its dual-phase architecture, which tactically combines unsupervised anomaly detection with supervised outcome forecasting. This model successfully detected individuals displaying unusual behavioral and psychological patterns, instances that are frequently overlooked in entirely supervised models. Eliminating or marking these anomalies decreased noise in the training data and enhanced the Random Forest classifier's performance. This discovery verifies that hybrid learning can tackle one of the significant issues in mental health analytics: the existence of rare yet clinically vital outliers. Conventional models often perform poorly in these situations, particularly when trained on unbalanced data or when outliers distort the training sample distribution.

Conclusively, from a practical viewpoint, DeepStressScan offers a clear and scalable solution that can aid mental health experts in the early detection of high-risk individuals. The enhanced recall and F1-scores demonstrate the system’s capacity to identify instances that require more careful consideration, which is crucial in mental health assessments where missed cases can lead to significant real-world repercussions. Additionally, the analysis of feature importance provides insight into the factors influencing stress predictions. Attributes such as sleep length, social engagement, and screen use are both quantifiable and adjustable, rendering them practical objectives for intervention and digital health coaching.

Appendices

Data Sources:

1. Data Set of Depression (DASS 21) and Emotional Reactivity (ERS) - Indonesian. Published: 26 April 2024. Available at https://data.mendeley.com/datasets/sk27t97ymb/1a

2. The RHMCD-20 datasets for Depression and Mental Health Data Analysis with Machine Learning. Published: 8 January 2024. Available at https://data.mendeley.com/datasets/pxjmjyfdh2/2

Additional Information

Author Contributions

All authors have reviewed the final version to be published and agreed to be accountable for all aspects of the work.

Concept and design: Femi Johnson, Adebayo Abayomi-Alli, Olugbenga S. Olukumoro , Oladapo Elugbadebo, Dada O. Aborisade

Acquisition, analysis, or interpretation of data: Femi Johnson, Adebayo Abayomi-Alli, Olugbenga S. Olukumoro , Dada O. Aborisade

Drafting of the manuscript: Femi Johnson, Adebayo Abayomi-Alli, Olugbenga S. Olukumoro , Oladapo Elugbadebo

Critical review of the manuscript for important intellectual content: Femi Johnson, Oladapo Elugbadebo, Dada O. Aborisade

Supervision: Femi Johnson, Adebayo Abayomi-Alli, Olugbenga S. Olukumoro , Oladapo Elugbadebo, Dada O. Aborisade

Disclosures

Human subjects: All authors have confirmed that this study did not involve human participants or tissue. **Animal subjects:** All authors have confirmed that this study did not involve animal subjects or tissue. **Conflicts of interest:** In compliance with the ICMJE uniform disclosure form, all authors declare the following: **Payment/services info:** All authors have declared that no financial support was received from any organization for the submitted work. **Financial relationships:** All authors have declared that they have no financial relationships at present or within the previous three years with any organizations that might have an interest in the submitted work. **Other relationships:** All authors have declared that there are no other relationships or activities that could appear to have influenced the submitted work.

References

1. Seitz-Holland J, Mulsant BH, Reynolds III CF, et al.: Major depression, physical health and molecular senescence markers abnormalities. Nature Mental Health. 2023, 1:200-209. 10.1038/s44220-023-00033-z
2. Johnson FT, Onashoga A, Ibharalu T, Opakunle V, Adenusi C: WORK-PERF: An intelligent predictive model for work performance rating. Proceedings of Third Emerging Trends and Technologies on Intelligent Systems. ETTIS 2023. Lecture Notes in Networks and Systems. Noor A, Saroha K, Pricop E, Sen A, Trivedi G (ed): Springer, Singapore; 2023. 730:11-20. 10.1007/978-981-99-3963-3_2
3. Obermeyer Z, Powers B, Vogeli C, Mullainathan S: Dissecting racial bias in an algorithm used to manage the health of populations. Science. 2019, 366:447-453. 10.1126/science.aax2342
4. Johnson FT: Modeling stealing madness ratings using artificial intelligence technique. COJ Robotics & Artificial Intelligence. 2024, 4:1-2. 10.31031/cojra.2024.04.000581
5. Sarda A, Munuswamy S, Sarda S, Subramanian V: Using passive smartphone sensing for improved risk stratification of patients with depression and diabetes: Cross-sectional observational study. JMIR mHealth and uHealth. 2019, 7:e11041. 10.2196/11041
6. Garcia-Ceja E, Riegler M, Nordgreen T, Jakobsen P, Oedegaard KJ, Tørresen J: Mental health monitoring with multimodal sensing and machine learning: A survey. Pervasive and Mobile Computing. 2018, 51:1-26. 10.1016/j.pmcj.2018.09.003
7. Fauzi MA, Yang B, Blobel B: Comparative analysis between individual, centralized, and federated learning for smartwatch based stress detection. Journal of Personalized Medicine. 2022, 12:1584. 10.3390/jpm12101584
8. Al Masud GH, Shanto RI, Sakin I, Kabir MR: Effective depression detection and interpretation: Integrating machine learning, deep learning, language models, and explainable AI. Array. 2025, 25:100375. 10.1016/j.array.2025.100375
9. Alakananda K, Ananth Prabhu G, Souza MD, Vidya VV: A systematic review on leveraging machine learning and deep learning for early mental health depression detection on social media platforms. South Eastern European Journal of Public Health. 2025, XXVI:4570-4592.
10. Birjali M, Kasri M, Beni-Hssane A: A comprehensive survey on sentiment analysis: Approaches, challenges and trends. Knowledge-Based Systems. 2021, 226:107134. 10.1016/j.knosys.2021.107134
11. Lodha P, Pandya A: Editorial: Ethical dilemmas of digitalisation of mental health. Frontiers in Human Dynamics. 2025, 7:1654355. 10.3389/fhumd.2025.1654355
12. Bhattacharya S, Prusty S, Pande SP, Gulhane M, Lavate SH, Rakesh N, Veerasamy S: Integration of multimodal imaging data with machine learning for improved diagnosis and prognosis in neuroimaging. Frontiers in Human Neuroscience. 2025, 19:1552178. 10.3389/fnhum.2025.1552178
13. Shatte ABR, Hutchinson DM, Teague SJ: Machine learning in mental health: a scoping review of methods and applications. Psychological Medicine. 2019, 49:1426-1448. 10.1017/s0033291719000151
14. Nykoniuk M, Basystiuk O, Shakhovska N, Melnykova N: Multimodal data fusion for depression detection

approach. Computation. 2025, 13:9. 10.3390/computation13010009

15. Koutsouleris N, Dwyer DB, Degenhardt F, et al.: Multimodal machine learning workflows for prediction of psychosis in patients with clinical high-risk syndromes and recent-onset depression. JAMA Psychiatry. 2021, 78:195-209. 10.1001/jamapsychiatry.2020.3604
16. Lattie EG, Adkins EC, Winquist N, Stiles-Shields C, Wafford QE, Graham AK: Digital mental health interventions for depression, anxiety, and enhancement of psychological well-being among college students: Systematic review. Journal of Medical Internet Research. 2019, 21:e12869. 10.2196/12869
17. Adler DA, Stamatis CA, Meyerhoff J, et al.: Measuring algorithmic bias to analyze the reliability of AI tools that predict depression risk using smartphone sensed-behavioral data. npj Mental Health Research. 2024, 3:57. 10.1038/s44184-024-00057-y
18. Rezaie Z, Banad Y: Machine learning applications in Alzheimer's disease research: a comprehensive analysis of data sources, methodologies, and insights. International Journal of Data Science and Analytics. 2025, 20:3169-3203. 10.1007/s41060-024-00651-5
19. Sharma SK, Alutaibi AI, Khan AR, Tejani GG, Ahmad F, Mousavirad SJ: Early detection of mental health disorders using machine learning models using behavioral and voice data analysis. Scientific Reports. 2025, 15:16518. 10.1038/s41598-025-00386-8
20. Schölkopf B, Platt JC, Shawe-Taylor J, Smola AJ, Williamson RC: Estimating the support of a high-dimensional distribution. Neural Computation. 2001, 13:1443-1471. 10.1162/089976601750264965
21. Tabassum M, Mahmood S, Bukhari A, Alshemaimri B, Daud A, Khalique F: Anomaly-based threat detection in smart health using machine learning. BMC Medical Informatics and Decision Making. 2024, 24:347. 10.1186/s12911-024-02760-4
22. Jaiswal A, Babu AR, Zadeh MZ, Banerjee D, Makedon F: A survey on contrastive self-supervised learning. Technologies. 2021, 9:2. 10.3390/technologies9010002
23. Srividya M, Mohanavalli S, Bhalaji N: Behavioral modeling for mental health using machine learning algorithms. Journal of Medical Systems. 2018, 42:88. 10.1007/s10916-018-0934-5
24. Rana N, Marwaha H: Auto encoder-guided feature extraction for pneumonia identification from chest X-ray images. E3S Web of Conferences. 2024, 556:01011. 10.1051/e3sconf/202455601011
25. Al-Abadi AAJ, Mohamed MB, Fakhfakh A: Enhanced random forest classifier with K-means clustering (ERF-KMC) for detecting and preventing distributed-denial-of-service and man-in-the-middle attacks in Internet-of-Medical-Things networks. Computers. 2023, 12:262. 10.3390/computers12120262
26. Lim JY, Lim KM, Lee CP, Tan YX: SCL: Self-supervised contrastive learning for few-shot image classification. Neural Networks. 2023, 165:19-30. 10.1016/j.neunet.2023.05.037
27. Vaswani A, Shazeer N, Parmar N, et al.: Attention is all you need. NIPS'17: Proceedings of the 31st International Conference on Neural Information Processing Systems. 2017, 6000-6010.
28. Sano A, Taylor S, McHill AW, Phillips AJ, Barger LK, Klerman E, Picard R: Identifying objective physiological markers and modifiable behaviors for self-reported stress and mental health status using wearable sensors and mobile phones: observational study. Journal of Medical Internet Research. 2018, 20:e210. 10.2196/jmir.9410
29. Pinge A, Gad V, Jaisighani D, Ghosh S, Sen S: Detection and monitoring of stress using wearables: a systematic review. Frontiers in Computer Science. 2024, 6:1478851. 10.3389/fcomp.2024.1478851
30. Pedregosa F, Varoquaux G, Gramfort A, et al.: Scikit-learn: Machine learning in Python. Journal of Machine Learning Research. 2011, 12:2825-2830.