International Scientific Journal of Engineering and Management (ISJEM) ISSN: 2583-6129
Volume: 04 Issue: 06 | June – 2025 DOI: 10.55041/ISJEM04551
An International Scholarly || Multidisciplinary || Open Access || Indexing in all major Database & Metadata

# Model for Recognizing and Understanding Neural AnxietyLevels

Ms.Chatura Narvekar
*Electronics and Communications Engineering*
*Agnel Institute of Technology and Design*
Goa,India
21ec04@aitdgoa.edu.in

Mr.Samir Mulla
*Electronics and Communications Engineering*
*Agnel Institute of Technology and Design*
Goa,India
21ec26@aitdgoa.edu.in

Mr.Shabi Gadekar
*Electronics and Communications Engineering*
*Agnel Institute of Technology and Design*
Goa,India
21ec28@aitdgoa.edu.in

Mr.Sudin Kerkar
*Electronics and Communications Engineering*
*Agnel Institute of Technology and Design*
Goa,India
21ec36@aitdgoa.edu.in

Mr.Mrunal Sawant
*Electronics and Communications Engineering*
*Agnel Institute of Technology and Design*
Goa,India
ms@aitdgoa.edu.in

Ms.Needhi Gaonkar
*Electronics and Communications Engineering*
*Agnel Institute of Technology and Design*
Goa,India
ng@aitdgoa.edu.in

***Abstract*—The project is an example of a stress detection system that uses a number of sensors to track physiological signals such body temperature, respiration rate, acceleration, EMG, ECG, and EDA. The device uses a unique analysis model to classify stress levels as low, medium, or high.A software platform processes the data gathered from sensors, such as the Seeed Studio Grove GSR Sensor, Spinkky MPU-6050 Module, Heart BioAmp, Muscle BioAmp Patch v0.2, and temperature sensor. The platform offers tailored stress management advice based on the user's stress level. The goal of this effective and user-friendly technology is to promote mental health and enable real-time stress monitoring.**

## I. Introduction

Stress has become an unavoidable aspect of modern life and exerts a profound impact on physical and mental well-being. It manifests itself in various forms, such as physical stress caused by fatigue or illness and psychological stress driven by emotional or cognitive factors such as fear, frustration, or worry. Although some stress can be beneficial in small doses, chronic or elevated levels lead to adverse effects, including anxiety, depression, sleep disorders, heart disease, and memory problems.

In today's fast-paced world, people often lack real-time awareness of their stress levels, making it difficult to address the problem proactive. India, in particular, faces alarmingly high rates of stress-related issues among young adults aged 15 to 29, highlighting the need for innovative solutions [**?**].Traditional stress management tools are largely reactive, focusing on addressing symptoms rather than preventing escalation.

Advancements in wearable technology and sensor systems provide an opportunity to bridge this gap. Real-time monitoring of physiological parameters, such as acceleration, body temperature, respiratory rate, electrodermal activity (EDA), electrocardiogram (ECG), and electromyography (EMG), allows for early detection of stress. This paper proposes a web application that leverages these parameters to analyze stress levels, categorizing them as low, medium, or high. The following sections of this paper discuss the methodology for developing the application, relevant machine learning models, implementation details, and the ability of the system to provide actionable feedback to users. By proactive stress management, this research aims to contribute to a practical and accessible solution to stress management.

## II. Literature Review

Recent studies have explored the use of machine learning techniques for detecting mental stress through physiological signals. A comprehensive review [2] highlighted the effectiveness of algorithms such as Support Vector Machines (SVM), Neural Networks, and Deep Learning in classifying stress levels using EEG, ECG, GSR, skin temperature, and heart rate variability. The integration of multimodal data sources, including facial expressions and speech, has been shown to improve accuracy, although ethical concerns, privacy issues, and hardware integration challenges persist. Another study [3] focused on wearable sensors for the detection of stress stress in free-living environusinging physiological signals such as ECG, skin temperature, and skin conductance. Features such as heart rate variability and phasic/tonic skin conductance components were linked to stress responses. Machine learning models like Random Forest and XGBoost achieved high accuracy, with pre-processing techniques like SMOTE effectively addressing data imbalances.

Further research has delved into fatigue detection using machine learning, examining physiological signals such as EMG, ECG, and EDA, combined with behavioral data such as gait and posture [4]. These studies emphasized the need for robust sensor calibration, consideration of environmental factors, and integration of multiple data sources for improved accuracy. The WESAD data set has been pivotal in multimodal stress detection research [5], using physiological characteristics from EDA, EMG, ECG, respiration rate, and temperature for stress classification. Both machine learning (Random Forest, SVM) and deep learning (Artificial Neural Networks) models were validated through Leave-One-Subject-Out (LOSO) cross-validation, demonstrating robust, subject-independent performance.

Personalized stress detection models have gained traction, particularly those that use EEG and ECG signals to create

 

International Scientific Journal of Engineering and Management (ISJEM) ISSN: 2583-6129
Volume: 04 Issue: 06 | June – 2025 DOI: 10.55041/ISJEM04551
An International Scholarly || Multidisciplinary || Open Access || Indexing in all major Database & Metadata

patient-specific models that significantly outperform general ones [6]. This personalized approach underscores the importance of accounting for interindividual variability in physiological responses. Furthermore,ensemble learning models trained on synthesized datasets combining SWELL, WESAD, and NEURO have demonstrated strong generalization capabilities in stress prediction, achieving accuracies up to 85 percent [7]. Innovations in wearable technology, such as the development of microfluidic electrochemical sensors for the noninvasive monitoring of oxidative stress biomarkers in sweat [8], and wrist-worn devices that detect stress-related behaviors via motion sensors [9], showcase practical applications of machine learning in real-time, non-invasive stress detection systems.

## III. Methodology

### A. *Dataset Overview*

The WESAD (Wearable Stress and Affect Detection) dataset, a widely used multimodal physiological dataset, forms the foundation for this research on stress detection. The data set includes data from 15 participants, each exposed to various conditions designed to elicit specific physiological responses. These conditions are as follows:

Baseline: Participants are in a neutral, relaxed state, providing a reference for normal physiological activity. Key indicators include stable heart rate, low electrodermal activity (EDA), and relaxed respiration.

Stress: Participants undergo stress-inducing tasks such as public speaking or timed problem-solving. Physiological markers include elevated heart rate, increased EDA, and rapid respiration.

Amusement: Positive emotional stimuli, like watching humorous videos, evoke moderate physiological arousal without stress. Indicators include slightly elevated heart rate and EDA.

Meditation 1 and 2: Guided meditation sessions promote relaxation, leading to decreased heart rate, lower EDA, and slow, deep breathing.

### B. *Dataset Preparation*

The WESAD dataset, consisting of physiological data from 15 subjects, was utilized for stress classification. It includes multimodal signals such as ECG, EDA, EMG, respiration rate, body temperature, and acceleration, recorded under three stress levels: Low, Medium, and High.

To ensure a well-balanced dataset, 5,000 samples per stress level were extracted from each subject, leading to a final dataset of 225,000 samples (75,000 per class). This balanced representation minimizes bias, prevents overfitting, and enhances the model's ability to generalize across diverse physiological patterns, ensuring robust stress classification.

### C. *Data Preprocessing*

To enhance data quality and ensure consistency across physiological signals, preprocessing was applied to the WESAD dataset before model training. Given the multimodal nature of the data, preprocessing involved signal synchronization, noise removal, missing data handling, normalization, and feature extraction.

All signals were resampled to a uniform frequency to maintain temporal alignment. Filtering techniques were applied to remove noise from ECG, EDA, EMG, and respiration signals, ensuring reliable feature extraction. Missing values were handled using interpolation and median imputation, minimizing data loss. Z-score normalization was performed to standardize the features, preventing scale variations from affecting model performance.

This structured preprocessing approach ensures data integrity, enhances model interpretability, and optimizes classification performance.

### D. *Sensor Data Conversion*

To convert the raw sensor values into SI units, each channel has to transformed based on the formulas given below:

1) ECG (mV): $\left(\frac{\text{signal}}{2^{16}} - 0.5\right) \cdot 3$
2) EDA (μS): $\left(\frac{\text{signal}}{2^{16}} \cdot 3\right) / 0.12$
3) EMG (mV): $\left(\frac{\text{signal}}{2^{16}} - 0.5\right) \cdot 3$
4) TEMP (°C): $\frac{\text{signal} \cdot 3}{2^{16} - 1}$
5) XYZ Acceleration (g): $\left(\frac{\text{signal} - 28000}{38000 - 28000}\right) \cdot 2 - 1$
6) Respiration (%): $\left(\frac{\text{signal}}{2^{16}} - 0.5\right) \cdot 100$

### E. *Correlation with stress levels*

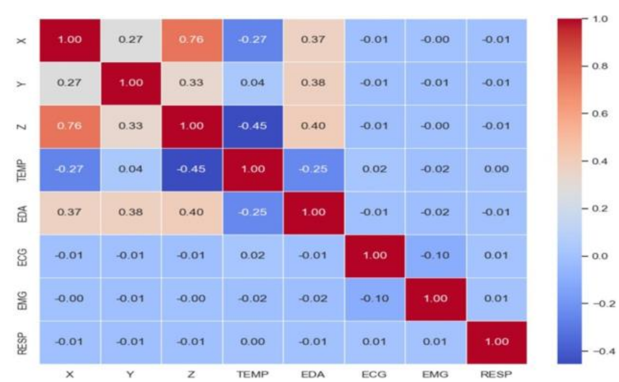


Fig. 1. Correlation matrix of numerical features

The heatmap reveals key relationships among physiological signals for stress classification. Strong correlations (red) indicate synchronized variations, while negative correlations (blue) highlight inverse trends.

The Z-axis acceleration strongly correlates with the X-axis (0.76), suggesting motion influence on stress responses. EDA moderately correlates with acceleration, linking skin conductance to movement. Temperature negatively correlates with Z (-0.45), reflecting thermoregulatory effects.

ECG, EMG, and RESP show minimal correlation with other features, contributing independently. This analysis refines feature selection, reducing redundancy and enhancing classification accuracy.

 

International Scientific Journal of Engineering and Management (ISJEM) ISSN: 2583-6129
Volume: 04 Issue: 06 | June – 2025 DOI: 10.55041/ISJEM04551
An International Scholarly || Multidisciplinary || Open Access || Indexing in all major Database & Metadata

### *F. Distribution of Features Across Stress Levels*

Understanding the distribution of physiological signals across stress levels is critical for developing robust machine learning models in stress classification. Violin plots provide an effective visualization of data density, variability, and feature separability, making them ideal for stress detection analysis.

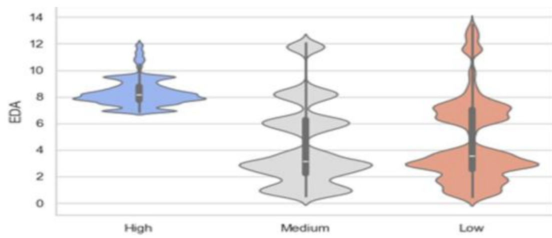


Fig. 2. EDA Distribution Across Stress Levels

Primary stress biomarker:EDA values increase significantly with rising stress levels, confirming its strong correlation with autonomic nervous system activation. High stress levels exhibit a wider distribution, reflecting individual variability in stress response. EDA serves as the most critical feature for stress classification in machine learning models.

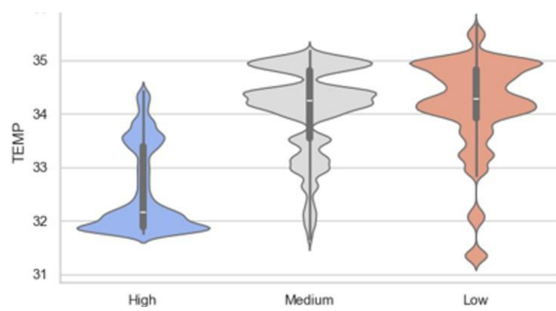


Fig. 3. TEMP Distribution Across Stress Levels

Inverse correlation with stress: Higher stress levels lead to lower skin temperature, likely due to sympathetic nervous system-induced vasoconstriction. Well-separated distributions between Low and High stress support its use as a secondary stress marker.

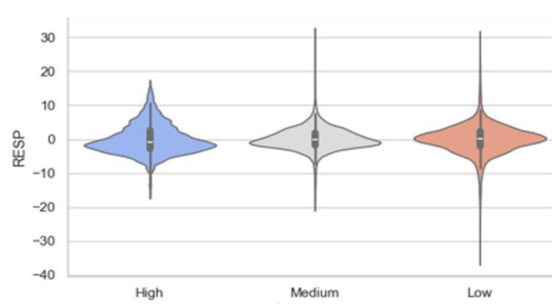


Fig. 4. RESP Distribution Across Stress Levels

Irregular breathing patterns under High stress, indicated by greater distribution spread. More stable breathing in Low and Medium stress states, reinforcing RESP as a secondary feature for classification.

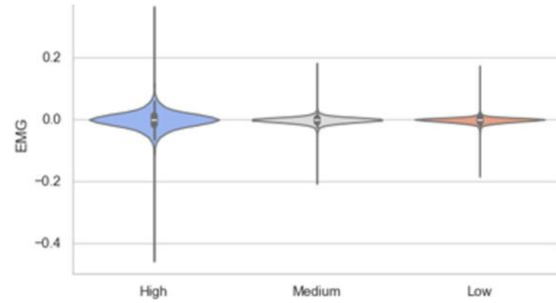


Fig. 5. EMG Distribution Across Stress Levels

Increased muscle activation in High stress, demonstrating stress-induced muscular tension. Medium and Low stress levels exhibit lower EMG values, supporting its role as a stress response indicator.

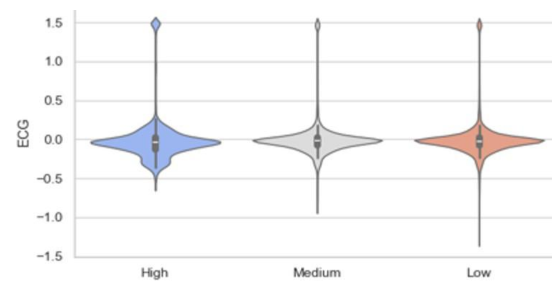


Fig. 6. ECG Distribution Across Stress Levels

ECG raw values show minimal separation across stress levels, indicating it may require derived features such as Heart Rate Variability (HRV) for effective classification.

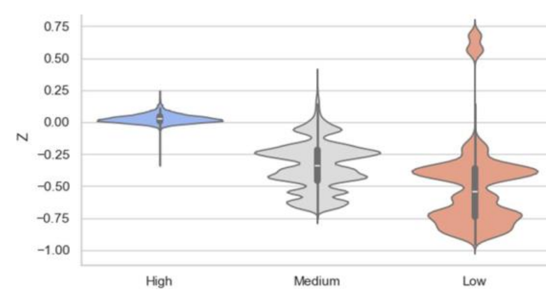


Fig. 7. ACC Distribution Across Stress Levels

Accelerometer data shows increased movement intensity in High stress states, with the Z-axis exhibiting the most variation due to postural adjustments and involuntary movements. X and Y axes remain relatively stable, indicating limited lateral or forward-backward motion under stress.

International Scientific Journal of Engineering and Management (ISJEM) ISSN: 2583-6129
Volume: 04 Issue: 06 | June – 2025 DOI: 10.55041/ISJEM04551
An International Scholarly || Multidisciplinary || Open Access || Indexing in all major Database & Metadata

## IV. Machine Learning Model

Machine learning techniques are extensively utilized in physiological signal classification due to their ability to capture intricate patterns and variations. This study implements multiple ML models, including Decision Trees, Random Forest, Support Vector Machines (SVM), K-Nearest Neighbors (KNN), Extra Trees, AdaBoost, and Gradient Boosting, to classify stress levels based on physiological data. Among these, Random Forest emerged as the most effective model due to its superior accuracy and robustness.

Random Forest is a robust ensemble learning algorithm that enhances classification accuracy by aggregating predictions from multiple decision trees. Each tree is trained on a random subset of the data, and final classification is determined through majority voting, improving generalization, and reducing overfitting. In this study, Random Forest was used for stress classification (Low, Medium, High) based on physiological signals, including electrodermal activity (EDA), temperature (TEMP), respiration (RESP), electrocardiogram (ECG), electromyography (EMG) and accelerometer (ACC) data.

### A. Working Mechanism

1. Bootstrap Aggregation (Bagging): Random Forest employs bootstrap sampling, where each decision tree is trained on a random subset of the dataset drawn with replacement. This process enhances diversity among trees, leading to a reduction in variance and improved generalization.

2. Random Feature Selection: At each node split, a random subset of features is selected rather than evaluating all available features. This prevents over-reliance on dominant features, ensuring that trees remain decorrelated and independent, thereby enhancing model robustness.

3. Majority Voting for Classification: Each decision tree in the forest independently predicts a stress level (Low, Medium, or High). The final classification decision is determined through majority voting, where the most frequently predicted class among all trees is chosen as the output.

4. Parallelization and Scalability: Since decision trees are trained independently, Random Forest supports parallel processing, significantly reducing computation time. This property makes it ideal for real-time stress classification in wearable and biomedical applications.

### B. Mathematical Representation

Given a dataset $D$ with $N$ training samples:

1) Create $k$ decision trees, where each tree $T_i$ is trained on a **random bootstrap sample** $D_i$ from $D$.
2) At each node, the best feature is chosen from a **random subset** $F_s \subset F$, where $F$ is the full feature set.
3) The final predicted class $\hat{Y}$ is determined by majority voting:

$$\hat{Y} = \text{mode}\{T_1(X), T_2(X), \ldots, T_k(X)\} \quad (1)$$

where $T_i(X)$ represents the prediction of the $i$-th tree for input $X$, and *mode* represents the majority vote across all decision trees.

### C. Advantages

• **High Accuracy:** Random Forest reduces overfitting by averaging the predictions of multiple trees, resulting in higher accuracy and generalization.
• **Robustness to Noise:** It is less sensitive to noisy data and outliers, which is beneficial when working with physiological signals that may have inherent variability.
• **Handles Mixed Data Types:** It can process both numerical and categorical data effectively, making it suitable for this project's diverse physiological inputs.
• **Scalability:** Random Forest works well with large datasets and high-dimensional data, which is useful when handling multiple physiological signals.

### D. Model Training and Evaluation

The Random Forest algorithm can be implemented using the RandomForestClassifier from the sklearn.ensemble library in Python.It requires the tuning of several hyperparameters to optimize performance such as

- **n_estimators** = 300: Number of trees in the forest.
- **criterion** = "gini": Uses the Gini impurity for splitting nodes.
- **random_state** = 3: Ensures reproducibility of results.
- **max_samples** = 0.5: Uses 50% of samples for training each tree.
- **max_features** = 0.75: Considers 75% of features when looking for the best split.
- **max_depth** = 25: Limits the depth of trees to prevent overfitting.
- **bootstrap** = True: Enables bootstrapping for sampling data.

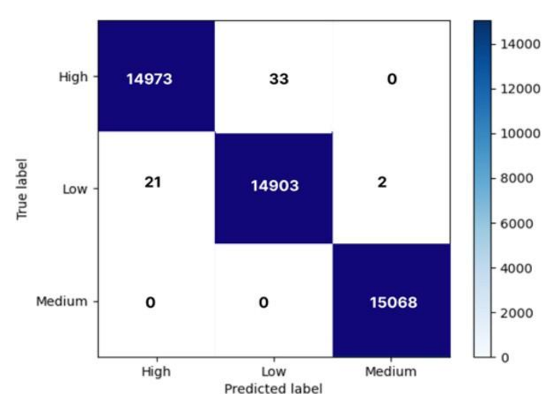


Fig. 8. Confusion Matrix for Random Forest

ISSN: 2583-6129
DOI: 10.55041/ISJEM04551

• Excellent classification across all stress levels:
o 33 instances of ”High” stress were misclassified as ”Low.”
o 21 instances of ”Low” stress were misclassified as ”High.”
o Only 2 instances of ”Low” stress were misclassified as ”Medium.”
• Random Forest is a reliable choice for high accuracy and robustness, particularly when interpretability is not a primary concern.

*E. Accuracy and Precision Comparison*

| ML Algorithms | Train Acc(%) | Test Acc(%) | Acc(%) | Precision (%) | | |
|---|---|---|---|---|---|---|
| | | | | LOW | MEDIUM | HIGH |
| Decision Tree | 98.81 | 98.75 | 98.75 | 99.99 | 97.95 | 98.31 |
| KNN | 99.02 | 98.69 | 98.69 | 99.25 | 98.4 | 98.42 |
| Random Forest | 99.99 | 99.87 | 99.87 | 99.85 | 99.77 | 99.98 |
| Extra Tree | 99.88 | 99.85 | 99.85 | 99.74 | 99.81 | 99.99 |
| Adaboost | 71.13 | 71.28 | 71.28 | 56.02 | 66.91 | 96.24 |
| Gradient Boosting | 99.79 | 99.77 | 99.77 | 99.66 | 99.67 | 99.99 |
| SVM | 99.33 | 99.23 | 99.23 | 98.69 | 99.05 | 99.96 |

## V. Hardware Design

The proposed system utilizes multiple biosensors and an inertial measurement unit (IMU) interfaced with an Arduino UNO for the acquisition of physiological signals and further processing.The key hardware components include:
1) **Heart BioAmp Candy:** Records electrocardiogram (ECG) signals for cardiac monitoring and extraction of respiration rate.
2) **Muscle BioAmp Patchy:** Measures electromyographic signals to analyze muscle activity.
3) **MPU6050 :** Provides motion tracking using accelerometer and body temperature data.
4) **Galvanic Skin Response:** Monitors skin conductance variations to assess physiological stress levels.

## VI. Web Application

The proposed web application is developed using Flask, HTML, and CSS, integrating a machine learning model (stress-ml-model.pkl) to analyze user inputs and generate personalized stress assessments. Additionally, an AI-powered chatbot, built with Automation Anywhere AI, enhances user interaction by providing real-time assistance, answering queries, and offering stress management recommendations.

The system follows a structured workflow, beginning with the homepage (home.html), where users are introduced to the platform. They then proceed to a questionnaire module (/basic question), where preliminary stress-related data is collected, followed by the detailed input module (/inputs), which gathers additional parameters for a more precise evaluation. The Flask backend processes this data using the pre-trained ML model, which predicts stress levels based on predefined parameters. The results are then translated into personalized feedback, delivered through an interactive user interface with visual aids such as stress-indicating images (breathing high.jpg, diet low.jpg), helping users understand their stress levels and recommended coping strategies.

To further improve user experience, the AI-integrated chatbot enables real-time engagement, guiding users through the platform and providing stress management insights. This combination of machine learning and AI-driven automation ensures an efficient, interactive, and accessible stress assessment tool, offering data-driven insights and promoting effective stress management practices.