iJRASET
International Journal For Research in
Applied Science and Engineering Technology

INTERNATIONAL JOURNAL
FOR RESEARCH
IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 13 Issue: IV Month of publication: April 2025

DOI: https://doi.org/10.22214/ijraset.2025.68711

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 13 Issue IV Apr 2025- Available at www.ijraset.com

Mental Stress Prediction Using Machine Learning

Tejal Purani[1], Nimesh Vaidya[2], Dr. Vijay Gadhavi[3]
[1]PG Scholar – Faculty of Engineering, Computer Engineering Department Swaminarayan University,
[2]Assistant Professor & HOD-Faculty of Engineering, Computer Engineering Department Swaminarayan University
[3]Associate Professor & Dean –Faculty of Engineering (I/C), Computer Engineering Department Swaminarayan University

Abstract: In today's fast-paced world, mental stress has become a significant issue, contributing to numerous health issues and lower productivity. Early detection and prediction of stress levels can help individuals take timely preventive measures. Analyzing physiological, behavioral, and questionnaire-based data, this project presents a machine learning-based method for predicting mental stress. The system leverages various machine learning algorithms, including Support Vector Machines (SVM), Random Forest, and Logistic Regression, to classify stress levels into different categories such as low, moderate, and high. Heart rate, sleep patterns, physical activity, and responses to psychological assessment scales (such as the DASS-21 or Perceived Stress Scale) are among the data features used for prediction. The dataset is preprocessed for noise reduction, normalization, and missing value handling. Feature selection techniques such as PCA and correlation analysis are applied to enhance model performance. In order to guarantee the models' robustness and generalizability, k-fold cross-validation is used to train and validate them. The Random Forest algorithm is a promising tool for real-time mental health monitoring because it predicts stress levels with the highest accuracy, as demonstrated by experimental results. This study demonstrates the potential for continuous stress monitoring and mental health support by combining machine learning with wearable sensors and mobile apps.
Keywords: Machine Learning (ML), Supervised Learning, Behavioral Data

I. INTRODUCTION

Stress is a physiological response to the emotional, mental, or physical challenges that become a serious problem affecting people of different life situations, and age groups in day-to-day life[2] . This study developed an activity-aware mental stress detection system using ECG, GSR, and accelerometer data from 20 participants. By incorporating activity data, it addressed the issue of physical activity masking mental stress responses. With a 10-fold cross-validation, the system was able to classify mental stress with an accuracy of 92.4% and classify between subjects with an accuracy of 80.9%, indicating potential for continuous, real-world stress monitoring.

A. Different Kinds of Stress

1) Mental stress which is short term stress and does not cause extensive damage. It is easy to detect and treatable. Mental stress is also known as acute stress. Acute Stress is short-lived. It can be beneficial and create motivation. For example, when a deadline is approaching, stress may help you to focus and complete your task before the deadline [1].
2) Episodic acute stress which makes people anxious. Acute stress that is suffered too frequently is called episodic stress [1]. This type of stress is usually seen in people who make self-inflicted, unrealistic or unreasonable demands [1]
3) Chronic stress can be detrimental to both physical and mental health. Chronic stress also plays a role in mental illnesses such as generalized anxiety disorder and depression. Chronic stress is difficult to manage because it cannot be measured in a consistent and timely way [1].

The most commonly used physiological parameters of stress are as follows [1]:

- Electrocardiograph
- Blood Pressure
- Galvanic skin response
- Electromyogram
- Skin temperature
- Body temperature
- Glucose level
- Pupil diameter

a) Electrocardiograph: The most commonly used stress marker parameters derived from the electrocardiograph are the heart rate and Heart rate variability (HRV) [1].
b) Blood pressure: When the body is under stress, it responds by producing a surge of stress hormone causing an increase in blood pressure and heart to beat faster. Individuals who underwent a stressful task had a late recovery of blood pressure making it a good indicator of stress even after some time.
c) Galvanic skin response: Using changes in skin conductivity. During Stress, resistance of skin drops due to increased secretion in sweating glands [1].
d) Electromyogram: measuring electrical activity of the muscles. Stress causes differences in the contraction of the muscles which can be used to identify stress [4].
e) Skin temperature: Changes in temperature of the skin are related to the stress level [4].
f) Body temperature: when we feel cold or chilly, the body wants to increase its body temperature because it is dropping toward the low end within the normal range. But that change is a minor one, and not more than plus or minus 1 F[10].

II. DEFINITION OF THE PROBLEM

A key challenge in detecting stress levels using machine learning techniques involves feature extraction, selection, feature combination, and improving performance metrics from EEG signals. To address these issues, We propose a hybrid feature dataset that enhances accuracy by incorporating both time and frequency domains, overcoming the limitations of current methods. Traditional approaches, such as subjective measures like questionnaires, are frequently used for stress detection but fail to accurately capture intrinsic mental states. These methods often lack the precision required for clinical settings. In contrast, proposed solution uses objective measures, specifically physiological signals like EEG, for stress detection. Based on the above motivating elements the problem statements can be defined as follows: “To design and develop an efficient machine learning model capable of accurately detecting stress levels by utilizing EEG signals and advanced machine learning techniques.”

III. SCOPE OF THE WORK: [5]

The study targets on detecting stress levels utilizing a hybrid dataset and hybrid classifier, employing machine learning techniques and feature selection methods used with EEG signals. The scope of the work is as follows:
1. Data Acquisition and Preprocessing: This process involves applying signal processing techniques to ensure the quality and reliability of the dataset.
2. Feature Extraction and Selection: To extract the features need to apply different techniques and to enhance the performance of the classification model, the most appropriate features will be identified through feature selection approaches. 3. Hybrid model: The study focus to prepare hybrid models for hybrid features & hybrid classifier by applying different ensemble methods. 4. Evaluations and comparison of results: The performance of the hybrid model will be evaluated using measures such as F1-score, recall, accuracy, and precision. 5. Applications and Implications: The developed system aims to support real-time stress monitoring and has potential applications in the medical field, workplace environments, and individual stress management.

IV. PROPOSED WORKFLOW TO DETECT STRESS LEVEL THROUGH MACHINE LEARNING TECHNIQUES

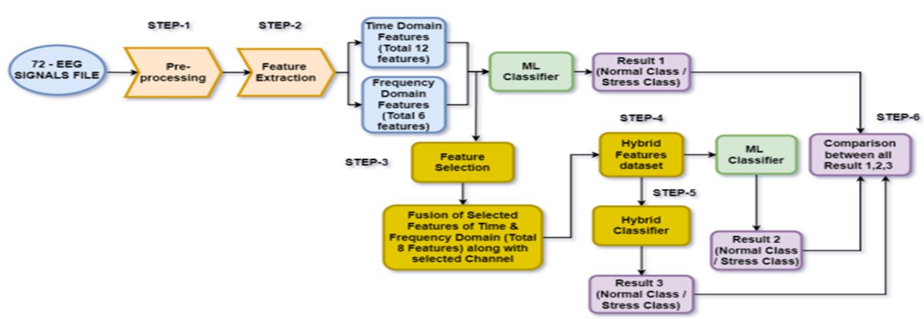

Figure 1: Schematic view of proposed stress level detection through machine learning

Figure 1 shows the Schematic view of proposed stress level detection through machine learning algorithms. It will be divided into the below steps.

1) Step 1: Pre-processing In the preprocessing phase, raw data (EEG signals) in EDF file format is processed and stored to assure its quality and relevance for further analysis. EEG signals are filtered using the Finite Impulse Response (FIR) method, which successfully removes the noise and artifacts while preserving the data's important features. Epoch segmentation is used to split the continuous EEG data into epochs after filtering. The combined effects of these preprocessing procedures make reliable feature extraction and classification possible in the stress detection framework.

Preprocessing performed two steps on EEG signals.
FIR – finite impulse response method: FIR filters are designed to reduce noise and unwanted frequency data from EEG signals while retaining important brain oscillation patterns including beta, theta, delta, and alpha waves. A principal merit of these filters lies in their stability, as they are fundamentally devoid of feedback mechanisms, thereby ensuring the absence of distortion or instability within the processed signal.
Epoch Segmentation: After applying the FIR method, the continuously filtered EEG data is divided into time windows, known as epochs. Each epoch represents a duration of 5 seconds, with a 1-second overlap between consecutive epochs.

2) Step 2: Feature Extraction In the feature extraction phase, both time domain and frequency domain features are extracted from the EEG signals. Time domain features to focus on the mean, standard deviation (std), variance, skewness, kurtosis, etc. frequency domain features obtaining using power spectrum density (PSD) through the welch method, which analyzes the band frequency features. These features provide vital information to detect stress levels in the EEG-based stress detection framework.
3) Step 3: Feature Selection In the feature selection phase, a feature importance ranking method is applied to identify the most distinct features from both the time and frequency domains.
4) Step 4: Hybrid features dataset In the hybrid features dataset phase, the selected features from both domains are fused after applying the feature selection method, resulting in a newly prepared hybrid features dataset.
5) Step 5: Hybrid Classifier In the hybrid classifier phase, the hybrid feature dataset is utilized as the input for constructing a hybrid classifier using the stacking method. This approach involves employing base learner classifiers alongside a meta-classifier to improve the model's predictive accuracy. The hybrid classifier effectively leverages the complementary strengths of different algorithms, enhancing its ability to handle complex EEG data. This results in a more reliable and robust detection of stress levels, ensuring high accuracy in the stress classification process.
6) Step 6: Comparision of Results In this phase, the results of the proposed model are compared with those of existing methods. This comparison highlights the hybrid classifier's gains in performance, accuracy, and reliability when it comes to detecting the stress level from EEG signals, showing the proposed method's efficiency over traditional approaches.

V. PREPARE A PROPOSED HYBRID FEATURES DATASET:[8]

Figure 2 illustrates the procedural framework for the compilation of the suggested hybrid features dataset. This process will be elucidated comprehensively through the proposed algorithm.

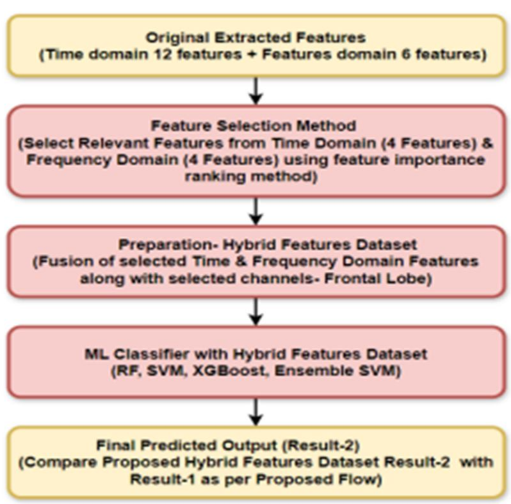

Figure 2: Workflow to prepare proposed hybrid features dataset

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 13 Issue IV Apr 2025- Available at www.ijraset.com

Algorithm : Feature Importance Ranking Method for Preparing Proposed Hybrid Features Dataset Input: Original Extracted Features (Considering both Time and Frequency Domain) from EEG Signals Initialize Model: Random Forest - machine learning model that provides feature importance scores

Data Preparation:

- Load and preprocess the dataset to ensure it is clean and ready for training.
- Split the dataset into training and testing sets.

Train the Model: Train the Random Forest model using the training dataset.

Extract Feature Importance:

- After training, extract the feature importance scores from the model.
- These scores indicate the contribution of each feature to the model's predictive power.

Rank Features:

- Rank all features based on their importance scores in descending order.
- Higher scores indicate more important features.

Select Top Features: Select the top 4 features based on their importance scores.

Output Selected Features: Output the selected top 4 features consider as a selected features to prepare Hybrid Dataset.

VI. PREPARE A PROPOSED HYBRID META CLASSIFIER (HMC):[7]

Figure 3 illustrates the procedural framework for the development of the proposed hybrid classifier. This will be comprehensively elucidated through the proposed algorithm. This process necessitates the identification of both the base learner and the meta-model to establish a hybrid classifier methodology. It is anticipated that this approach will augment evaluation metrics such as precision, recall, accuracy, and F1-score.

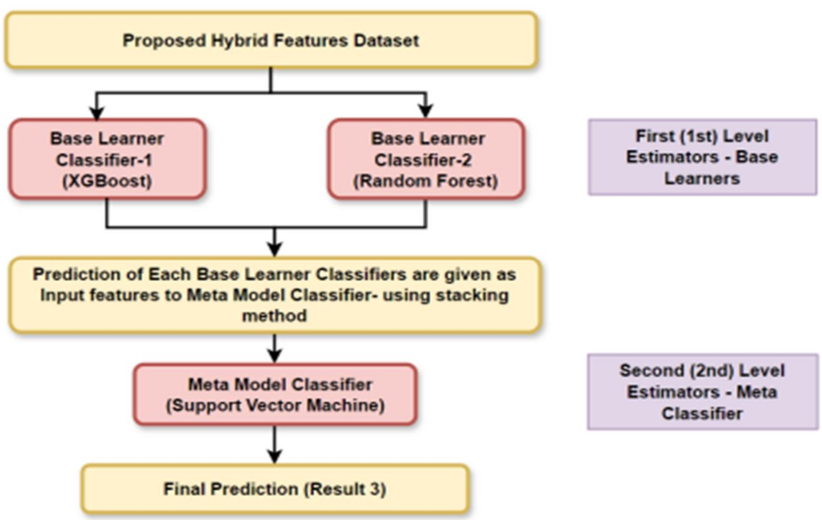

Figure 3: Workflow to prepare proposed hybrid classifier

Algorithm : Hybrid Stacking- Based Classification Algorithm.

Input: Hybrid features dataset

Split Hybrid Dataset: Hybrid dataset split into training and testing sets.

Define Base Learners: consider Random Forest (RF) and XGBoost (Extreme Gradient Boosting) as a base Learner model.

Apply Hybrid approach through Stacking Method: [6]

- Train RF on the training set
- Train XGBoost on the training set
- Use both RF and XGBoost to predict on the testing set
- Create a new dataset (stacked features) with predictions from RF and XGBoos as new features

Define Meta Models: Support Vector Machine (SVM) consider as a meta model Train Meta Models on Stacked Features: Train SVM on the stacked features (predictions from RF and XGBoost base learners models)
Final Predictions: Use the trained SVM to predict on the testing set.
Output: Final Hybrid Classification Model with Stacking

VII. CONCLUSION

This study shows machine learning, especially [best-performing model], can accurately predict mental stress from physiological and behavioral data. This non-invasive method allows for early stress detection using wearable sensors and other data sources. Future research should focus on larger, diverse datasets and real-time integration into apps for improved stress management.

EXAMPLE OF LIST OF REFERENCES

[1] Dr Des McLernon , Dr L Mhamdi : Analysis and Processing physiological data from a watch-like device to detect stress pattern, The University of Leeds, August 2015.
[2] Feng-Tso Sun , Cynthia Kuo , Heng-Tze Cheng , Senaka Buthpitiya , Patricia Collins : Activity Aware Mental Stress Detection Using Physiological Sensors, Institute for Computer Sciences, Social Informatics and Telecommunications Engineering 2012.
[3] J. Minguillon, E. Perez, M. A. Lopez-Gordo, F. Pelayo, and M. J. Sanchez-Carrion, "Portable system for real-time detection of stress level," Sensors, vol. 18, no. 8, p. 2504, Aug. 2018, doi: 10.3390/s18082504.
[4] Kalpesh Patil , Manisha Singh , Garima Singh :"Mental Stress Evalution using Heart rate variability analysis : A review" , ISSN volume 2 , April-2015
[5] F. Akhtar, M. B. B. Heyat, J. P. Li, P. K. Patel, and R. Guragai, "Role of machine learning in human stress: A review," in 17th International Computer Conference on Wavelet Active Media Technology and Information Processing, 2020, pp. 170-174, doi: 10.1109/ICCWAMTIP51612.2020.9317396.
[6] A. Nirabi, F. A. Rahman, M. H. Habaebi, K. A. Sidek, and S. Yusoff, "Machine learning-based stress level detection from EEG signals," in IEEE 7th International Conference on Smart Instrumentation, Measurement and Applications (ICSIMA), 2021, pp. 53-58, doi: 10.1109/ICSIMA50015.2021.9526333.
[7] O. AlShorman, M. Masadeh, M. B. B. Heyat, F. Akhtar, H. Almahasneh, G. M. Ashraf, and A. Alexiou, "Frontal lobe real-time EEG analysis using machine learning techniques for mental stress detection," J. Integr. Neurosci., vol. 21, no. 1, pp. 1-11, Jan. 2022, doi: 10.31083/j.jin2101020.
[8] L. D. Sharma, V. K. Bohat, M. Habib, A. M. Al-Zoubi, H. Faris, and I. Aljarah, "Evolutionary inspired approach for mental stress detection using EEG signal," Expert Systems with Applications, vol. 197, p. 116634, Feb. 2022, doi: 10.1016/j.eswa.2022.116634.
[9] J. Agrawal, M. Gupta, and H. Garg, "Early stress detection and analysis using EEG signals in machine learning framework," IOP Conference Series: Material Science and Engineering, vol.1116, no.1, p. 012134, 2021, doi: 10.1088/1757-899X/1116/1/012134.