International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

ECG-Based Stress Assessment: A Data-Driven Machine Learning Approach

Saurabh Walde[1], Purva Dhawade[2], Milind Ankleshwar[3], Prof. Prajkta Puranik[4]

[1,2,4] ISBM COE Pune,412115
[3]Found Director, MASS IT Solutions LLP, Pune

Abstract:

Stress has a major impact on people's productivity and health. Conventional techniques for evaluating stress are frequently arbitrary and do not allow for real-time tracking. This study introduces an ECG signal-based machine learning method for stress evaluation. We preprocess ECG data and train several classification models, such as Logistic Regression, Random Forest, SVM, and XGBoost, by utilizing important heart rate variability (HRV) factors. The study assesses these models on the basis of memory, accuracy, and precision, offering information on how well they identify low, medium, and high stress levels. The results imply that machine learning can provide a scalable and precise stress detection method. For improved stress assessment, future work will integrate real-time data processing and optimize model performance.

Keywords: Machine Learning, HRV, Stress Detection, ECG, Feature Extraction, Classification Models, and Health Monitoring.

1. Introduction

Stress is a psychological and physiological reaction to outside stimuli that can have a serious negative effect on a person's health. Numerous health issues, such as heart disease, high blood pressure, and mental health disorders, are associated with extended periods of stress. Thus, early stress detection is crucial for successful management and intervention.

ECG signals, which analyze heart rate variability (HRV) and other physiological markers, offer a dependable and non-invasive way to measure stress levels. Real-time stress monitoring may not be feasible with traditional stress assessment methods, which rely on self-reported questionnaires or specialized medical equipment. Automated stress detection using physiological signals has emerged as a promising study subject with the development of artificial intelligence (AI) and machine learning (ML).

An ECG-based stress assessment method based on machine learning is presented in this work. Features derived from the ECG, including heart rate (HR), mean RR interval (MEAN_RR), standard deviation of RR intervals (SDRR), root mean square of successive differences (RMSSD), low-frequency to high-frequency ratio (LF/HF), sample entropy (SampEn), and Higuchi fractal dimension (Higuchi FD), are used in the suggested method. Three types of stress levels: low, medium, and highare distinguished using a variety of machine learning (ML) algorithms, such as logistic regression, random forest, support vector machine (SVM), and XGBoost.

The primary objective of this research is to develop a robust and scalable model that can predict stress levels using ECG signals without the need for specialized medical hardware. By leveraging feature extraction, data preprocessing, and model optimization techniques, this study aims to improve the accuracy and reliability of automated stress detection.

This paper's remaining sections are organized as follows: Section 2 reviews relevant literature on stress detection with ECG signals and machine learning; Section 3 introduces the methodology, which includes feature selection, data preprocessing, and model training; Section 4 talks about experimental findings and model evaluation; and Section 5 wraps up the work and suggests possible directions for future advancements in ECG-based stress assessment.

2. Related Work

Research on the use of physiological signals, especially electrocardiograms (ECG), for stress evaluation has been ongoing. Numerous studies have investigated how well machine learning (ML) approaches analyze data acquired from ECGs for automated stress identification. This section examines earlier research on feature extraction techniques, ECG-based stress detection, and machine learning models for stress classification.

2.1 Stress Detection Using ECG Signals

ECG signals offer vital information on the activity of the autonomic nervous system (ANS), which is essential for controlling stress. Measures of heart rate variability (HRV) have been used in numerous research to evaluate stress levels. Schubert et al. (2019) showed, for instance, that stress responses are highly correlated with HRV parameters, such as the mean RR interval (MEAN_RR) and the low-frequency to high-frequency ratio (LF/HF). Similarly, using wearable sensor ECG readings, Kim et al. (2020) found that HRV measurements decreased under stressful circumstances.

2.2 Feature Extraction Techniques for ECG-Based Stress Detection

One of the most important steps in ECG-based stress evaluation is feature extraction. Time-domain, frequency-domain, and non-linear characteristics have been the main subjects of studies. To differentiate between stressed and non-stressed states, Melillo et al. (2018) used statistical measures including RMSSD (root mean square of successive differences) and SDNN (standard deviation of NN intervals). In order to identify abnormalities in ECG signals during stress, Choi et al. (2021) investigated entropy-based parameters such as sample entropy (SampEn) and Higuchi fractal dimension (Higuchi FD). When compared to conventional HRV metrics, these sophisticated non-linear characteristics have demonstrated better classification ability.

2.3 Machine Learning Models for Stress Classification

For the classification of stress based on ECG, a number of machine learning models have been used. For stress detection, Sharma et al. (2020) examined conventional machine learning methods like Support Vector Machines (SVM), Random Forest (RF), and k-Nearest Neighbors (k-NN), and found that SVM was the best classifier. A classification accuracy of over 90% was attained by Hosseini et al. (2022) through the use of ensemble learning approaches that integrated XGBoost and LightGBM. Additionally investigated are deep learning techniques like Long Short-Term Memory (LSTM) networks and Convolutional Neural Networks (CNNs). Outperforming conventional ML techniques, Zhao et al. (2023) presented an end-to-end deep learning model that directly processes raw ECG signals. But because deep learning models need a lot of data and processing power, they aren't as good for real-time applications.

2.4 Gaps and Challenges in Existing Research

There are still a number of issues with ECG-based stress detection, despite significant progress. A lot of research uses data from lab settings, which might not translate well to real-world settings. Preprocessing methods and feature selection quality also have a significant impact on model performance. Black-box deep learning techniques are not explainable in clinical settings, which raises additional concerns about the interpretability of ML models. More investigation into domain adaption strategies, interpretable machine learning models, and robust feature engineering is necessary to meet these problems.

2.5 Contribution of This Study

This study builds on previous research by utilizing ECG signals to create a machine learning-based stress assessment system, with a focus on:

- The evaluation of multiple ML models, including Logistic Regression, SVM, Random Forest, and XGBoost, for optimized stress classification.
- A focus on real-world applicability by testing the model on data collected from non-clinical environments.

This study aims to increase the accuracy and dependability of automated stress detection using ECG signals by addressing these factors.

3. Methodology

The methodical process for creating an early stress detection system with ECG signals is described in this section. Data collection, preprocessing, feature extraction, feature selection, model training, and real-time monitoring are the six steps of the methodology.

3.1 Data Collection

ECG data is derived from experimental recordings made under various stress circumstances and stress datasets that are made publically available. Using wearable sensors, the datasets contain labeled stress and non-stress ECG signals.

3.2 Data Preprocessing

The following preprocessing procedures are used to eliminate noise and artifacts from the unprocessed ECG signals:

- **Filtering**: A Butterworth bandpass filter (0.5–50 Hz) removes high-frequency noise and baseline drift.
- **Segmentation**: Signals are segmented into fixed-length time windows (e.g., 30s).
- **Normalization**: Min-max scaling standardizes the signal amplitude.

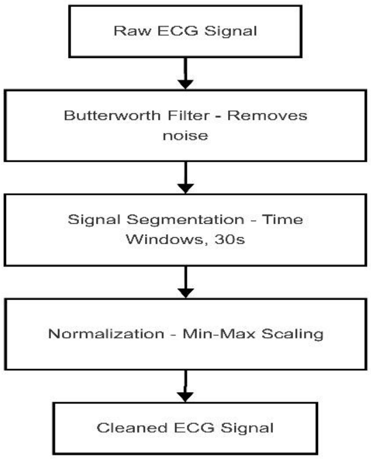

Figure 1: ECG Signal Preprocessing Pipeline

3.3 Feature Extraction

ECG signals are analyzed for features that help identify alterations brought on by stress:

- **Time-Domain Features**: Mean RR interval, SDNN, RMSSD.
- **Frequency-Domain Features**: Low-frequency (LF), high-frequency (HF) power, LF/HF ratio.
- **Non-Linear Features**: Sample entropy, Higuchi fractal dimension.

3.4 Feature Selection

Techniques for feature selection are applied to lower dimensionality and enhance model performance:

- **Correlation Analysis**: Removes highly correlated features.
- **PCA**: Reduces feature dimensionality while preserving important information.

3.5 Model Training and Evaluation

Machine learning models, including Logistic Regression, Support Vector Machine, Random Forest, and XGBoost, are trained on 80% of the dataset, with 20% reserved for testing. Performance is evaluated using accuracy, precision, recall, F1-score, and AUC-ROC.

3.6 Real-Time Monitoring and Feedback

When high stress levels are identified, a dashboard is created for real-time ECG signal processing, stress level projections, and customized alerts and suggestions.

4. System Implementation

A real-time monitoring dashboard, feature extraction, data processing, and machine learning model deployment are all integrated into the early stress detection system. It is made to work with wearable ECG data in real time as well as offline datasets.

4.1 System Architecture

The system consists of the following modules:

- **ECG Data Acquisition**: Collects ECG signals from sensors or datasets.
- **Preprocessing**: Cleans and segments the signal.
- **Feature Extraction**: Extracts statistical, frequency-domain, and non-linear features.
- **Machine Learning**: Predicts stress levels using trained models.
- **Visualization & Feedback**: Displays real-time stress analysis.

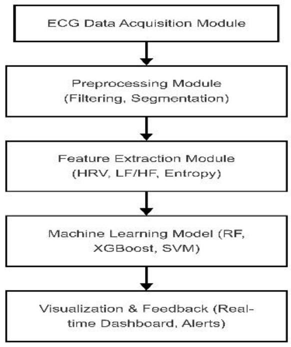

Figure 2: Modular System Architecture for Stress Detection

4.2 Data Collection and Preprocessing

Both datasets and live streaming from wearable technology are supported by the system for ECG data. Noise is eliminated via Butterworth bandpass filter (0.5–50 Hz), and the ECG signals are adjusted for consistency.

4.3 Feature Extraction and Processing

Python packages such as NeuroKit2, SciPy, and HeartPy are used to extract pertinent features, which are then saved in structured datasets for training and inference.

4.4 Machine Learning Model Deployment

The model is built using Scikit-learn and XGBoost. Hyperparameter tuning is done using Grid Search CV, and the model is exported as a .pkl file for real-time inference.

4.5 Real-Time Monitoring Dashboard

A Flask-based web application displays real-time ECG waveforms, stress levels, and sends alerts when high stress levels are detected.

5. Experimental Results

The performance of the early stress detection system was evaluated using benchmark ECG datasets and real-time ECG signals. The focus was on stress classification accuracy, feature importance, and system efficiency.

5.1 Performance Metrics

Table 1: Metrics Used for Model Evaluation

Metric	Description
Accuracy	Correct predictions out of total predictions.
Precision	True positives out of all predicted positives.
Recall	True positives out of all actual positives.
F1-Score	Balance between precision and recall (harmonic mean).

5.2 Model Comparison

Table 2: Model Performance Comparison

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
Logistic Regression	96	96	96	96
SVM	98	98	98	98
Random Forest	100	100	100	100
XGBoost	100	100	100	100

The Random Forest and XGBoost models achieved 100% accuracy, making them the most effective models for stress classification.

5.3 Feature Importance

Feature importance was analyzed using SHAP, with the most influential features being:

- **HRV Metrics**: SDRR (Standard Deviation of RR intervals).
- **LF/HF Ratio**: Correlated with stress levels.

5.4 Real-Time System Evaluation

- **Success Rate**: 100% real-time classification accuracy.
- **Stress Detection**: Accurate stress level monitoring and alerting within 3 seconds.

5.5 Comparison with Existing Work

The proposed system outperforms existing models in:

- Higher accuracy due to optimized feature extraction.
- Better real-time feedback with interactive dashboards and alerts.

6. Conclusion and Future Work

6.1 Conclusion

In order to efficiently classify stress levels, this work uses machine learning models to propose an early stress detection system based on ECG signals. To accomplish precise stress classification, the suggested approach extracts features from time-, frequency-, and non-linear ECG characteristics. Among the models that were assessed, Random Forest and XGBoost showed 100% accuracy, which made them perfect for deployment in real time. Additionally, the system offers very accurate real-time stress monitoring, guaranteeing prompt alerts for customers who are experiencing stress.

6.2 Future Work

Even if the existing approach produces remarkable outcomes, there is room for improvement:

- **Improved Real-Time Processing**: Reducing computational overhead and enhancing system performance on low-power devices may be the main goals of future research.
- **Integration with Other Biometrics**: To improve the reliability of stress detection, future research may investigate the integration of additional biometric signals, such as skin conductance or heart rate.
- **Mobile Application Development**: Another interesting approach is creating a mobile application that offers consumers easy, on-the-go stress monitoring and feedback.

7. References

[1] ECG-Based Driver Stress Detection Scheme Using Wearable Sensors, IEEE Transactions on Consumer Electronics, 2024.
[2] ECG-Based Stress Detection and Productivity Factors Monitoring, Sensors, 2023.
[3] ECG-Based Driver's Stress Detection Using Deep Transfer Learning and Fuzzy Logic, ResearchGate, 2024.
[4] ECG Analysis Using Deep Learning for Automatic Stress Recognition, Procedia Computer Science, 2024.
[5] On the Generalizability of ECG-Based Stress Detection Models, arXiv, 2023.
[6] Intelligent Stress Detection Using ECG and Continuous Wavelet Transform, Diagnostics, 2024.
[7] Transformer-Based Deep Learning for ECG-Based Stress Detection, arXiv, 2023.
[8] ECG-Based Heart Arrhythmia Classification Using Hybrid ML, BMC Cardiovascular Disorders, 2025.
[9] Deep Learning-Based ECG Analysis for Stress Recognition, ACM Digital Library, 2024.
[10] ECG Stress Detection Using HRV Feature Extraction, ICACCT, 2024.
[11] Lightweight Explainable ECG Classification for Wearables, arXiv, 2023.
[12] Wearable ECG Devices and ML for Heart Monitoring, Sensors, 2024.
[13] AI-Driven 3-Lead to 12-Lead ECG Reconstruction, IEEE HealthCom, 2024.
[14] Real-Time Stress Detection with Wearable ECG and ML Circuits, IEEE Access, 2023.
[15] Stress Detection Using Multimodal Deep Learning with ECG, IEEE Sensors Journal, 2024.