**Optimization of Wearable Biosensor Data for Stress Classification Using Machine Learning and Explainable AI**

**ABSTRACT**

This work utilizes wearable devices for real-time stress detection and investigates the effectiveness of meditation audio in reducing stress levels after academic exposure. Physiological data, including Interbeat Interval (IBI)-derived Heart Rate Variability (HRV), Blood Volume Pulse (BVP), and electrodermal

activity (EDA), are collected during the Montreal Imaging Stress Task (MIST). The stress classification methodology employs an integrated approach using Genetic Algorithm and Mutual Information to reduce feature set redundancy. It further uses Bayesian optimization to fine-tune machine learning hyperparameters. The results indicate that the combination of EDA, BVP, and HRV achieves the highest classification accuracy of 98.28% and 97.02% using the Gradient Boosting (GB) algorithm for 2-level and 3-level stress classification. In contrast, EDA and HRV alone achieve a comparable accuracy of 97.07% and 95.23% for 2-level and 3-level stress classification, respectively. Furthermore, the SHAP Explainable AI (XAI) analysis confirms that HRV and EDA are the most significant features for stress classification. The study also finds evidence that listening to meditation audio reduces stress levels. These findings highlight the potential of wearable technology combined with machine learning for real-time stress monitoring and management in academic environments.

### Existing System

The existing system utilizes various machine learning algorithms, including Random Forest (RF), Decision Trees (DT), Support Vector Machine (SVM), K-Nearest Neighbors (KNN), Extreme Gradient Boosting (XGB), and Gradient Boosting (GB) for stress classification based on physiological signals like Electrodermal Activity (EDA), Blood Volume Pulse (BVP), and Heart Rate Variability (HRV). Among these, GB provides the highest classification accuracy at 98.28% for 2-level classification and 97.02% for 3-level classification, demonstrating superior performance over other models. This system also employs a hybrid feature selection approach, combining Genetic Algorithm (GA) and Mutual Information (MI), to reduce feature redundancy and optimize performance. Bayesian optimization is used for fine-tuning hyperparameters, further enhancing model accuracy and efficiency.

While the existing system demonstrates high classification accuracy, it has limitations. It relies heavily on pre-selected physiological signals, which may not capture all aspects of stress under different scenarios. Additionally, the system’s interpretability remains limited, even though the use of SHAP for explainability has improved insight into key features like HRV and EDA. Lastly, while effective in structured testing environments, this system may not generalize well in uncontrolled real-world settings, such as daily academic activities.

### Disadvantages

1. The system is optimized for controlled testing conditions, potentially reducing accuracy in varied real-world environments.
2. Relies mainly on HRV and EDA, potentially overlooking other influential stress indicators.
3. Although SHAP improves transparency, complex models like GB may still limit explainability for end-users.

### Proposed System

The proposed system aims to address these limitations by incorporating advanced deep learning models and expanding data inputs. Key algorithms include Convolutional Neural Networks (CNNs) for pattern recognition, Recurrent Neural Networks (RNNs) and Long Short-Term Memory (LSTM) networks for sequential data analysis, and Graph Neural Networks (GNNs) for capturing complex relationships between physiological and environmental factors. This system also integrates new data sources, such as ambient noise and temperature, alongside physiological signals, allowing for a comprehensive understanding of stress triggers in dynamic environments. The Temporal Fusion Transformer (TFT) is included to enhance temporal sensitivity, capturing patterns in stress response over time.

Furthermore, the proposed system utilizes AutoML techniques for efficient model selection and parameter optimization, improving adaptability across diverse settings. By incorporating more data sources and advanced models, this system enhances real-time stress detection accuracy and adaptability, providing a robust solution for varying academic scenarios.

### Advantages

1. Adapts to diverse environments with expanded data inputs, offering more reliable stress detection in daily academic settings.
2. Advanced models with explainability features enable more actionable insights for users and educators.
3. Incorporates environmental and physiological factors, enhancing sensitivity to real-time stress indicators across varied conditions.

**SYSTEM REQUIREMENTS**

➢ **H/W System Configuration:-**

➢ Processor - Pentium –IV

➢ RAM - 4 GB (min)

➢ Hard Disk - 20 GB

➢ Key Board - Standard Windows Keyboard

➢ Mouse - Two or Three Button Mouse

➢ Monitor - SVGA

**SOFTWARE REQUIREMENTS:**

* **Operating system :** Windows 7 Ultimate.
* **Coding Language :** Python.
* **Front-End :** Python.
* **Back-End :** Django-ORM
* **Designing :** Html, css, javascript.
* **Data Base :** MySQL (WAMP Server).