BRAIN. Broad Research in Artificial Intelligence and Neuroscience
e-ISSN: 2067-3957 | p-ISSN: 2068-0473
Covered in: Web of Science (ESCI); EBSCO; JERIH PLUS (hkdir.no); IndexCopernicus; Google Scholar; SHERPA/RoMEO; ArticleReach Direct; WorldCat; CrossRef; Peeref; Bridge of Knowledge (mostwiedzy.pl); abcdindex.com; Editage; Ingenta Connect Publication; OALib; scite.ai; Scholar9; Scientific and Technical Information Portal; FID Move; ADVANCED SCIENCES INDEX (European Science Evaluation Centre, neredataltics.org); ivySCI; exaly.com; Journal Selector Tool (letpub.com); Citefactor.org; fatcat!; ZDB catalogue; Catalogue SUDOC (abes.fr); OpenAlex; Wikidata; The ISSN Portal; Socolar; KVK-Volltitel (kit.edu)
2026, Volume 17, Issue 2, pages: 316-331
Submitted: January 24th, 2026 | *Accepted for publication:* March 23rd, 2026

Prediction of Stress-Induced Substance Use via a Multimodal Data Acquisition Framework and an Ensemble Machine Learning Model

Bijoy Chhetri
J C Bose University of Science and Technology, Faridabad, India.
bijoychhetri@gmail.com,
https://orcid.org/0000-0002-2513-2996

Lalit Mohan Goyal
J C Bose University of Science and Technology, Faridabad, India.
lalitgoyal78@gmail.com,
https://orcid.org/0000-0003-4618-0281

Mamta Mittal*
Delhi Skill and Entrepreneurship University, New Delhi, India.
mamta.mittal@dseu.ac.in,
https://orcid.org/0000-0003-0490-4413

Abstract: *The importance of understanding stress-related substance use episodes is a key research area, as it offers new insight into the association between stress and substance use behaviours. Existing methods are typically limited to analysing self-reported or isolated physiological signals, without providing real-time contextual analysis. This study proposes a multimodal approach for data acquisition and an ensemble machine learning technique to analyse how stress affects an individual's substance use behaviour. The dataset was acquired using wearable sensors and a psychological stress assessment questionnaire with a craving intensity scale. A dataset consisting of 1,325 instances was acquired from 53 voluntary participants in certified recovery environments in North-East India. Furthermore, a fusion-based Ensemble Random Forest Machine Learning (ERFML) model is proposed to analyse an integrated dataset of physiological, psychological, and craving features. It has been observed from the experiments that the proposed model has higher prediction accuracy (AUC=0.95) for stress-induced substance use. Furthermore, a positive correlation between stress and craving was identified (r=.73). Likewise, heart rate and electrodermal activity features reflect physiological imbalances that take place during stress(r=0.60). Approximately 70% of craving instances were observed to co-occur with elevated stress levels within a defined temporal window, where stress and craving measurements were aligned at the same or preceding time step. The model is further validated using Leave-One-Subject-Out cross-validation to ensure subject-independent generalisability. This fusion-based model provides enhanced prediction of addiction vulnerability prediction using wearable biosignals and psychological assessments to manage relapse and recovery during treatment.*

Keywords: *stress; machine learning; substance use; addiction; alcohol use disorder.*

How to cite: Chhetri, B., Goyal, L. M., & Mittal, M. (2026). Prediction of stress-induced substance use via a multimodal data acquisition framework and an ensemble machine learning model. *BRAIN. Broad Research in Artificial Intelligence and Neuroscience, 17*(2), 316-331. https://doi.org/10.70594/brain/17.2/19

* *Corresponding author*

1. Introduction

Stress is an interaction between psychological and physiological processes that adversely affects human health (Ovsiannikova et al., 2024). A variety of health issues are associated with chronic stress, such as cardiovascular disease, anxiety, depression, compromised immunity, substance abuse and addiction (Nikbakhtzadeh et al., 2023). Stress affects the Sympathetic-Adreno-Medullary (SAM), Hypothalamic-Pituitary-Adrenal (HPA), and the dopaminergic reward systems, which are involved in addictive behaviour (Duffing et al., 2014). Additionally, psychological stress decreases executive function, cognitive flexibility, and decision-making, which are established risk factors for substance abuse. Consequently, chronic substance use stimulates stress response systems to promote continuous hyperarousal, leading to dependence and increased relapse risk (Sinha, 2008). Numerous studies have found that stress is not just a consequence of addiction, but it is, in most cases, the underlying cause of its initiation, maintenance, and perpetuation (Sinha, 2024; Ruisoto & Contador, 2019; Schwabe et al., 2011). Also, repeated exposure to high levels of stress leads to maladaptive coping strategies, including alcohol and drug use (Koob et al., 2023; Koob & Volkow, 2016), leading to a cycle of stress-induced substance use and psychological disequilibrium.

Classical approaches to monitoring stress related to substance use mostly rely on self-reporting or physiological assessments (Mimura & Griffiths, 2003). These approaches can provide useful snapshots, but they do not capture the complex dynamics of stress and its interaction with body, mind, and intensity of urge for substance use. Recent developments in healthcare technology, including the emergence of wearable sensing technologies and Machine Learning (ML) algorithms, have opened the door to real-time, personalised healthcare, particularly in stress assessment (Vavrinsky et al., 2021). The wearable technology-based model continuously monitors physiological signals like galvanic skin responses (GSR), heart rate (HR), and body temperature using wearable sensors. The sensor responses are analysed to identify objective physiological stress using various classifiers, suggesting that substance use can be minimised if physical stress is regulated (Dankovich et al., 2024). On the other hand, subjective stress levels are measured via assessment tools prescribed to understand the gravity of stress that a particular person is subjected to during an altered situation. The Perceived Stress Scale (PSS) is one such assessment tool used to evaluate how different situations affect perceived stress and emotional responses. The items in this instrument assess feelings and thoughts experienced over the past month (Tavolacci et al., 2013). Such large-scale data from physiological, psychological, and contextual sources are analysed and applied in personalised interactive applications such as Fitbit, mHealth, and Arogya Setu (Balbim et al., 2021).

From the recent literature study, it is evident that there has been a significant amount of research in the field of stress and its impact on substance use. The preceding section briefly reviews the major contributions that have impacted the model development for the management of stress and its relation with substance use. A researcher in one of such studies (Al-Atawi et al., 2023) evaluated the potential for real-time stress detection using ML and wearable IoT sensors, integrating HR, GSR, and skin temperature to monitor stress states relevant to substance abuse. This study highlighted that despite the practical implementation on a prototype level, widespread integration of wearable biosensors into substance use treatment and management is still hindered. An integrated approach involving just-in-time interventions for stress and craving of substance use suggests some hope for generalisability and portability of wearable devices into the human body in the review of wearable-based closed-loop neuromodulation (Shrestha et al., 2023). In this work, it is reported that physiological stress indicators like heart rate variability (HRV), GSR, and skin temperature are strongly associated with stress-induced craving and substance use episodes. Likewise, the study (Bolpagni et al., 2024) found that across multiple wearable platforms for dynamic stress monitoring, HRV and electrodermal activity (EDA) remain the most sensitive biomarkers. These findings are further supported by another study in which HR and acceleration data served as predictor features; the models achieved up to 84.5% accuracy in detecting elevated stress levels, which are key components in predicting vulnerability to drug use among college students (Razavi et al., 2023). Additionally, Takano et al. (2023) reported that wearable biosignals combined with self-reported stress measures can predict consumption episodes in individuals with alcohol and methamphetamine use. It is also found in the literature that self-supervised learning on EDA data has drastically reduced the data-labelling burden, making personalised stress prediction more feasible when it is combined with validated psychological assessment scales (Kargarandehkordi et al., 2024). Another cohort study on students using wearables to assess stress found that it correlates strongly with periods of elevated substance use risk during cyclical phases of group stress (Neigel et al., 2024).

A benchmark lab-based experimental dataset on stress and its effect on individuals, the Wearable Stress and Affect Detection (WESAD) dataset (Schmidt et al., 2018), is one of a kind, considering research on stress detection and analysis of its effects. This dataset has been publicly accepted by the research community to derive insights that are always taken into consideration during model building for stress management. WESAD is a publicly available dataset for wearable stress and affect detection. It includes multiple features comprising physiological and motion data, recorded from both wrist- and chest-worn devices across 15 subjects during a laboratory study**.** It comprises data acquired from sensor modalities like blood volume pulse, electrocardiogram, EDA, electromyogram, respiration, body temperature, and three-axis acceleration. Keeping this dataset as a baseline, an existing study (Xefteris et al., 2023) has used IoT for mental health monitoring, which is a multimodal approach combining both physiological and psychological datasets. This study is crucial in detecting emotional states even while in substance use. There are some other evident studies that have proven to have an explicit link between real-time stress estimation and addiction risk prediction (Carter et al., 2023; Habelt et al., 2020; Jambhale et al., 2022; Sun et al., 2024). In these studies, multiple stress indicators are used in substance use prediction models to enable personalised prevention and relapse reduction interventions.

Existing literature suggests that there has been significant progress in the study of stress-related disorders and their consequences on substance use. The proposed study on the prediction of stress-induced substance usage in recovery environments using a fusion of a multimodal data acquisition framework and an ensemble ML model for early detection of vulnerability in addiction stages during the treatment and recovery sessions of the participants of the recovery environments. Substance is here broadly defined as alcohol, drugs of abuse, cannabis, and heroin. The proposed model consists of a data collection module for physiological stress using wearable sensors and PSS-based assessment for psychological stress indicators. Furthermore, craving intensity is measured for each participant. Participants were recruited from recovery settings such as rehabilitation centres, hospital outpatient departments (OPDs), and addiction medicine centres. The proposed model extracts the important features that are predictors of substance consumption linked with stress, and they are used further in addiction risk estimation in terms of the level of attention that is required during the treatment phase in the recovery environment.

The work contributes to the following:

- Integration of real-time stress level measurement with substance-use risk assessment in a proposed multimodal framework, addressing a critical gap in addiction research.
- Implementation of a prototype system using ESP32-based wearable sensors to monitor physiological stress indicators during cravings and withdrawals.
- Integration of PSS responses with physiological biosignals through ERFML models to enhance stress classification accuracy.

2. Materials and Methodology

Substance consumption has been a social menace globally, but how stress affects the consumption patterns has to be understood. In order to study the effect of stress on the cues of substance consumption, an integrated multimodal fusion-based ensemble ML framework has been proposed. The methodology integrates data collected from physiological signals using wearable sensors, psychological stress assessments through the PSS questionnaire, and the behavioural craving intensity scale into a unified decision-making system. The benchmark WESAD dataset, together with a real-time custom dataset, has been used to evaluate the feasibility of stress detection using wearable devices and to classify instances into attention classes. The attention classes A1 to A4 signify levels ranging from basic to critical clinical attention required during the recovery process.

2.1. Benchmark WESAD Dataset

WESAD is an established benchmark dataset for stress and affect recognition, featuring 15 subjects under stressful conditions. This dataset was acquired via synchronised ECGs, EDAs, accelerometers, and self-reported stress scales. This dataset represents a significant advancement over previous laboratory studies on stress and emotion by introducing three primary affective states, namely neutral, stress, and amusement. Further, this dataset contains self-reported perceived stress states of the participants. The features that are included in the dataset are heart rate variability, blood volume pulse, electrocardiogram signal, EDA, electromyogram, respiration, body temperature, and three-axis acceleration. In the benchmark dataset, the features have been identified to observe the physiological temperament of the body reacting towards stress, making it suitable for stress classification model development.

2.2. Experimental Setup to Acquire Real-time Dataset

An experimental setup has been formally designed to acquire a real-time dataset. For this purpose, a trained healthcare professional was involved in collecting the study's sample data in a recovery environment. It should be noted that the real-time component in this study refers to the real-time processing and inference component during acquisitions only. During data collection sessions, participants were in contexts such as withdrawal, counselling, post-therapy, or routine check-ups. During each session, participants were equipped with an ESP32-based wearable device integrated with physiological sensors in order to accumulate physical data. The device consists of MAX30102 for HR and SpO_2 monitoring, MLX90614 for skin temperature measurement, and a GSR sensor for EDA detection. In addition to this, participants are also administered PSS questionnaires. At each sampling instance, participants reported their substance craving intensity using a 5-point scale ranging from 0 to 4, where 0 indicates no craving, and 4 represents very high craving. Data collection was conducted at event-driven sampling points, including withdrawal episodes, counselling sessions, post-therapy periods, and routine monitoring intervals. This allows the dataset to capture real-world variations in physiological and psychological states of body and mind. A comprehensive real-time dataset was acquired through this multimodal approach, comprising 1,325 instances from 53 voluntary participants. The feature set considered during dataset preparation is represented in Table 1.

The proposed framework is based on well-established neurobiological mechanisms involved in stress processing, namely the SAM and HPA systems. The SAM system controls immediate autonomic responses, such as increased heart rate and EDA, which can be reliably measured using wearable devices like HR and GSR sensors. On the other hand, the HPA axis modulates slower hormonal responses, including the release of cortisol, impacting the cognitive appraisal of stress and subsequent behaviours, such as cravings. Skin temperature measures the vasoconstriction in the periphery mediated by the sympathetic nervous system and is partially affected by the HPA-mediated hormonal effects, but it is less sensitive than GSR and HR. Finally, SpO_2 acts as a physiological metric that remains stable and has little association with stress. Together, these parameters allow for the construction of a physiological model that accounts for SAM-mediated responses, psychological assessment driven by HPA-mediated cognitive appraisal, and craving via a behaviourally activated response mediated by dopaminergic reward pathways. The customised dataset from the participants also consists of categorical values of patients' stages at which the participants are undergoing the treatment, and session context during which session of recovery treatment their data has been collected. These features, along with the craving score, mainly derive the contextual information, and further, all the other information features are numerical in nature.

Table 1. Feature description of the customised dataset acquired from the participants in the recovery environment

Feature Name	**Data Type**	**Description**	**Unit/Range**	**Example**
Timestamp	Date Time	Time of data recording	ISO 8601	2025-05-10
Patient Stage	Categorical	Clinical stage in treatment	Beginner/Moderate/ Severe Dependence	Moderate
Heart Rate	Integer	Pulse rate (MAX30102)	BPM	76
SpO_2 Level	Integer	Oxygen saturation	%	98
Body Temperature	Float	Skin temperature (MLX90614)	°C	36.8
GSR Value	Integer	Skin conductance	μS	8.75
PSS Score	Integer	Total PSS questionnaire score	0–40	22
Psychological Stress	Categorical	PSS-derived category	Low/Medium/ High	Medium
Stress Probability	Float	Physio model confidence	0–1	0.82
Physio Stress	Integer	Derived physio probability	0/1	1
Craving Score	Integer	Willingness to use substance at sampling time	0–4	3
Session Context	Categorical	Collection context	Withdrawal/ Counselling/Post-therapy/ Routine	Withdrawal

After resampling and aligning sensor signals, biometric features were normalised for uniform scaling across modes. The timestamp variable indicates the specific moment in time when the physiologic and behavioral measures were recorded. In the present work, the dataset is considered an uncorrelated set of time-stamped instances that represent the moments of time when recordings take place in a controlled environment. Hence, timestamps serve for synchronisation purposes and do not imply the presence of any time-dependent processes. Even though temporal dependencies are not taken into account in the current task, the availability of the timestamp attribute allows future expansion. Further, the “Patient Stage” variable is indicative of the degree of severity of the substance abuse disorder and serves as a categorising label for the behaviour dimension of the model. It is assigned according to a clinical assessment that takes into account behavioural traits, frequency of substance abuse, and functional impairment during treatment. In the current research, patients were classified into three categories:

- Beginner: Intermittent or initial use of the drug with no significant functional impairment and dependence indicators.
- Moderate: Consistent use of substances with dependency tendencies, behavioural changes, and psychosocial impact.
- Severe Dependence: Compulsive consumption of drugs with pronounced symptoms of dependence and functional impairment with a high potential for relapse.

This classification is based on clinical opinion and can be considered as analogous to diagnostic tools such as the DSM-5 or ICD-11, particularly in respect to severity of dependence and functional impairment criteria. Nevertheless, an explicit scoring algorithm did not form part of this study, and patients’ actual condition has been used to inform the classification process.

The PSS scores have been subsequently categorised into Low, Medium, and High psychological stress levels based on established thresholds. The definition of the probability of stress is stated in the form of a numeric range of values between 0 and 1 (inclusive). This is a result of the probabilistic assessment from the model of the Random Forest (M1), calculated as a ratio of the number of decision trees in the forest identifying a certain example as “stress.”. Physio Stress refers to the dichotomous classification generated based on the probability of stress with a particular threshold level. This threshold level is 0.5, meaning that all probabilities above 0.5 are categorised as “stressed,” while those below 0.5 are classified as “not stressed.” These categories are further used to represent different levels of substance dependence severity. The entire dataset is pseudonymised to protect participant identity, and strict access control mechanisms have been implemented. Data transmission and storage are secured using encrypted channels, and the study is conducted following ethical clearance and informed consent from all participants.

2.3. Proposed Methodology

The proposed multimodal data acquisition framework to analyse the link between stress and substance use consists of a wearable device, a mobile application, and a cloud-based backend service. At the initial level, biosignals are collected using physiological sensors and processed to derive corresponding numerical values. These values are fed into an ML-based stress model that analyses the dataset and produces a probabilistic output of the model to classify stressed/not stressed. The probability values are derived from model confidence, calculated as the ratio of trees predicting a given class to the total number of trees in the ensemble. Furthermore, the psychological and craving intensity data are fed into a decision inference ML model to obtain PSS levels and craving scores. The decision fusion module is the final component of the system, classifying outputs into A1-A4 attention levels.

The clinical definitions of the attention levels (A1-A4) to clarify how a clinician would act upon each level has been elaborately represented in Table 2. In this module, features derived from the decision inference stage are fused and fed into an ML model for classification. It is important to note that the ML Stress Models (M1–M3) function as first-level learners, whereas the Decision Fusion Module acts as a second-level meta-classifier, forming a hierarchical ensemble learning architecture. The final deliverable of the proposed model is various levels of attention, i.e., Basic, Moderate, Elevated, and Critical attention that an individual is most likely to need while undergoing the treatment in a recovery environment.

The framework of the proposed fusion Ensemble Random Forest Machine Learning (ERFML) model is presented in Figure 1. For generalisability and balanced class representation, the dataset was stratified using a 70:30 train-test split with 5-fold cross-validation. For achieving evaluation independent of the subjects and avoiding any bias due to baseline learning of the model for individual participants, Leave-One-Subject-Out (LOSO) cross validation was used. In this technique, data of one subject was kept separate as the testing data, and the training data was obtained from other participants. This process was followed until each participant’s data was utilised once in the test phase, leading to 53 evaluation folds.

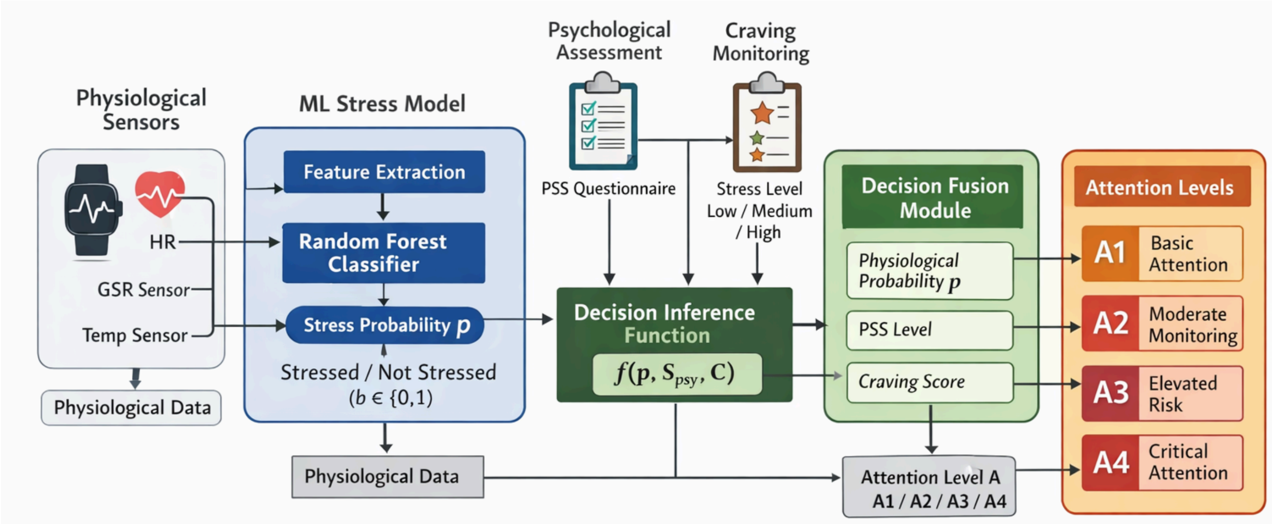

Figure 1. Proposed framework for Multimodal Data Acquisition Framework and an Ensemble Random Forest Machine Learning Model (ERFML)

The final output of the model is a mobile application programmed to automatically initiate intervention protocols based on predicted attention levels (A1–A4). In less risky cases (A1-A2), the application activates interventions through providing breathing and mindfulness instructions, motivational quotes, and cognitive reframing practices that focus on reducing stress. For higher risk cases (A2-A3), activation of interventions including cognitive behavioural therapy exercises, reminders of coping mechanisms, and monitoring techniques to manage cravings and emotions. In highly critical cases (A4), the mobile application automatically escalates to alerting clinical staff members, seeking assistance immediately, and accessing emergency counselling services.

Table 2. Clinical definition of attention levels and recommended clinical action

Attention Level	Clinical Definition	Risk Interpretation	Recommended Clinical Action
A1 (Basic Attention)	Elevated stress with minimal behavioural urge	Low immediate relapse risk	Routine monitoring, reassurance, preventive counselling, lifestyle guidance
A2 (Moderate Attention)	Emerging behavioural risk with moderate stress	Early warning stage	Brief intervention (SBIRT), stress management training, increased follow-up frequency
A3 (High Attention)	Significant stress with strong craving signals	High relapse risk	Structured therapy (CBT), targeted counselling, behaviour intervention, treatment plan adjustment
A4 (Critical Attention)	Concurrent high stress and high craving	Imminent relapse risk	Immediate intervention, intensive therapy, possible pharmacological support, supervised care

2.4. Working of Proposed ERFML Model

To effectively capture the heterogeneous nature of stress, psychological perception, and craving behaviour, a structured ensemble-based fusion model is proposed. Unlike conventional ML approaches that treat all features uniformly, the proposed model decomposes the problem into three domain-specific learning tasks and integrates them through a fusion ensemble framework. The physiological dataset acquired through wearable devices was validated by an expert physician for its viability; subsequently, it was manually classified into stressed and non-stressed classes based on clinically relevant signal features. After preprocessing, individual features are used to derive a composite physiological stress score, represented by the interpretable approximation in Equation (1):

$$Stress\ Score(P_s) = M1\ (X_{Ps})\ where\ X_{ps} = w_1 \cdot HR + w_2 \cdot GSR + w_3 \cdot Temp + b \quad (1)$$

where w_1, w_2, and w_3 are the learned feature weights, and b is the bias term, which is used to shift the decision boundary to get the maximum F1-score. The physiological stress score is classified using an ML classifier (M1), namely a Random Forest (RF) model. This RF model captures objective biometric stress signal features from wearable sensors.

Subjective psychological stress is assessed using the PSS administered through a mobile-based interface. With a predefined threshold of classification, the individual responder is classified into low stress (PSS <=13), medium stress (14< PSS <= 26), and considered high perceived stress (26< PSS <=40). A classifier model is trained on PSS-labelled data to dynamically categorise responses into Low, Medium, and High psychological stress levels. The recovery environments maintain PSS records of each attribute for every participant, enabling monitoring of intra-individual stress variations throughout the treatment process. This approach allows clinicians to track psychological stress trends across different treatment phases, such as withdrawal, counselling, post-therapy recovery, and normal routine check-ups. This classifier model (M2) learns cognitive stress perception patterns based on questionnaire responses as represented by equation (2).

$$Stress\ Score\ (P_{sh}) = M2\ (X_{psh}) \quad (2)$$

In addition to stress monitoring, participants report their craving intensity on a scale from 0 to 4, to introduce a behavioural dimension into relapse risk prediction. Craving intensity has been a strong indicator of potential substance-seeking behaviour. Patients are further categorised according to their clinical stage of care: Beginner, Moderate, and Severe Dependence. This classifier model captures behavioural intention and substance-seeking tendencies as per equation (3).

$$Craving\ Score\ (P_c) = M3\ (X_{pc}) \quad (3)$$

To design the final fusion model (M_{fusion}), outputs of the base models are combined to form a new feature vector as in equation (4). This hierarchical learning structure allows the model to learn interactions between different modalities, rather than raw feature-level correlations.

$$Z = M_{fusion}\{ P_s\ ,\ P_{sh}\ , P_c\} \quad (4)$$

The above equation combines body, mind, and the urge into a holistic approach of modelling all three datasets into a combined classification as one of the attention labels (A1-A4). The system integrates the information from multiple sources and derives a logical decision, often useful for the clinician. The algorithm for the classification of attention labels using the proposed ERFML model is presented below.

Algorithm for classification of attention labels using the proposed ERFML model

```
Input: HR, GSR, Temp, PSS attributes, Craving
Step 1: Train base models
    M1 = RF(HR, GSR, Temp), M2 = RF(PSS attributes), M3 = RF(Craving)
Step 2: Generate new features
    p1 = M1(x),    p2 = M2(x),    p3 = M3(x)
Step 3: Form a fusion vector
    Z = [p1, p2, p3]
Step 4: Train fusion model
    Mfusion = RF(Z)
Step 5: Predict    Output = Mfusion(Z)
Return: Attention Level
```

3. Experimental Results

This section highlights the major findings of the experimental analysis that has been performed on the benchmark WESAD dataset, as well as the real-time acquired dataset. An exploratory data analysis performed on the latter dataset reveals descriptive statistics of 1,325 instances. The statistics represent session-wise monitoring of participants undergoing treatment in a recovery environment. In order to assess the direct relationship among the quantitative variables, correlation analysis has been performed. Furthermore, trends and clusters were assessed using HexBin plots and box plots.

The descriptive statistics presented in Table 3, indicate that physiological signals are well distributed with HR and GSR. It suggests that there is high variability and a strong indication of stress. Overall, the population indicates that there is a moderate stress with a mean stress of ~19.45. During the session of data acquisition, the craving features have shown that mean craving (~1.94) shows wide behavioural variation, which acts as an essential feature for relapse modelling.

Table 3. Descriptive statistics of the customised dataset acquired from participants(N=53, n=1325)

Feature	Mean	Std Dev	Min	Median	Max
HR (BPM)	79.01	10.86	60	78	111
SpO_2 (%)	97.91	1.08	95	98	100
Temp (°C)	36.39	0.27	35.6	36.38	37.23
GSR (µS)	3.90	1.76	0.5	3.79	9.24
PSS Stress	19.45	9.66	5	17.59	39.92
Craving	1.94	1.20	0	2	4

As far as the correlation analysis is concerned, it is observed from the dataset that there is a strong correlation between PSS stress and craving ($r = 0.737$). Heart rate and EDA (~0.60) also indicate high potential to validate physiological sensing, whereas SPO_2 (~ -0.184) shows a weak negative relation that acts as a stabilising signal.
In order to measure the correlation between stress and craving, an associating criteria of time was defined. When craving occurs at time t, it is said to be associated with stress when the value of stress is above a threshold in a specific time interval of $[t-\Delta t,t]$. In this proposed experimental analysis, Δt represents one observation time frame. From this analysis, it has been found that approximately 70% of craving experiences have been noted to coincide with high levels of stress.

The HexBin plot shown in Figure 2(a) indicates a clear trend, whereby higher stress levels are predominantly associated with higher craving clusters.The HexBin visualisation reveals that higher stress levels tend to coincide with increased craving intensity, supporting the hypothesis of a stress–craving relationship.

Similarly, the box plot suggests that higher stress levels correspond to increased craving intensity. This is illustrated in Figure 2(b).

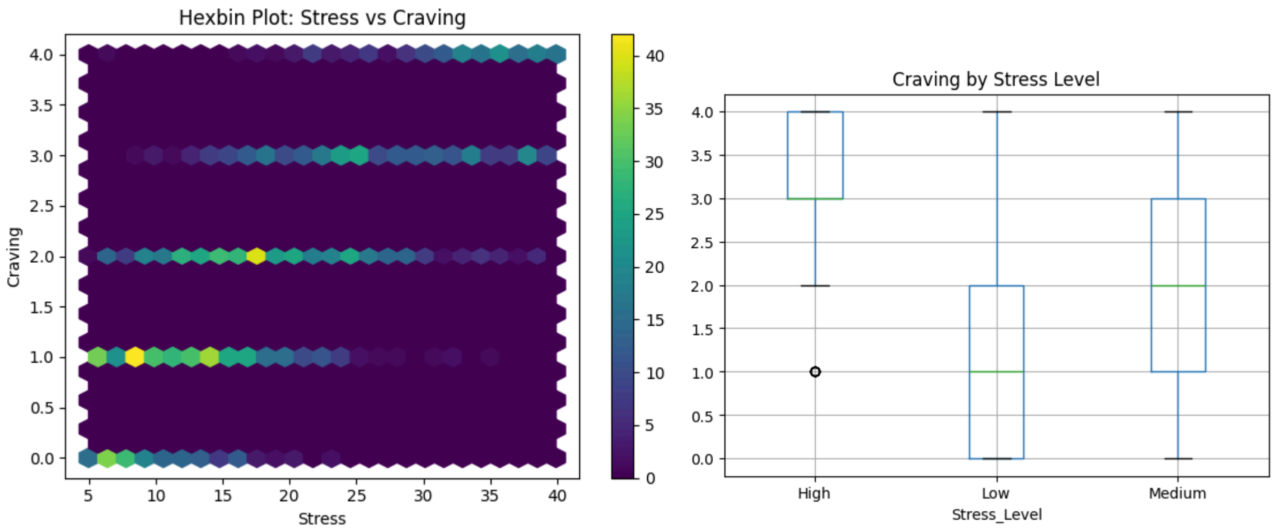

Figure 2. a) HexBin plot depicting how close stress appears over the craving cluster b) Box plot illustrating the distribution of craving scores across different stress levels

The performance of each ML classifier model has been assessed using standard performance metrics. The performance result is calculated for all the existing classifier models that are used and the proposed ERFML model.The evaluation metrics included accuracy, precision, recall, and F1 score. Furthermore, the area under the curve (AUC) derived from the receiver operating characteristic. In order to analyse the prediction accuracy, the confusion matrix was also plotted for the proposed model. Feature importance and model explainability were analysed using feature importance plots and SHapley Additive exPlanations (SHAP).

Results from both the WESAD benchmark dataset and the custom dataset are presented in Tables 4–6. From the experiment on both datasets, it is evident that the multimodal approach succeeded over the other in concluding that stress levels correlate with substance craving episodes.

The experimental results demonstrate consistent performance improvements of the proposed model over existing models on the real-time dataset. On the WESAD benchmark dataset, multiple classifier models were applied; however, the Random Forest (RF) model achieved the highest accuracy of 94%, confirming its robustness in capturing physiological stress patterns from wearable sensor data, as presented in Table 4.

Table 4. Performance of existing ML models on benchmark WESAD dataset

Model	Accuracy	Precision	Recall	F1-score
KNN	0.86	0.85	0.84	0.84
SVM	0.89	0.88	0.87	0.87
Neural Network	0.91	0.90	0.90	0.90
Random Forest	0.94	0.93	0.92	0.92

Similarly, in PSS-based psychological stress classification, the RF model outperformed other models with an accuracy of 93%, highlighting its effectiveness in modeling structured PSS-based features as shown in Table 5. As the best-performing base model, the RF classifier contributes to the improved performance of the proposed ERFML model. The hyperparameters considered for this model included number of trees (n_estimators),maximum depth (max_depth), minimum samples split (min_samples_split), minimum samples per leaf (min_samples_leaf), maximum features (max_features), bootstrap sampling, Gini impurity and random state. In the model, n_estimators = 100 was selected to ensure sufficient stability while maintaining computational efficiency. No limit was imposed on the maximum depth to allow the model to capture complex nonlinear relationships within the data. Bootstrapping was used to improve generalisability. The performance metrics, as depicted in Table 6, indicate that the model achieved a highest accuracy of 95% and an F1-score of 0.93. Moreover, the results highlight that the proposed fusion-based model has performed better in achieving the objective of the study by integrating physiological, psychological, and craving information.

Table 5. Performance of existing ML models on realtime dataset

Model	Accuracy	Precision	Recall	F1-score
SVM	0.90	0.89	0.89	0.89
KNN	0.87	0.86	0.85	0.85
Neural Network	0.91	0.90	0.90	0.90
Random Forest	0.93	0.92	0.91	0.91

From the experimental analysis, it has been observed that existing WESAD serves as a strong benchmark dataset for physiological stress detection and can be extended to various stress-related applications; however, it lacks integration of psychological assessment, which is critical for understanding perceived stress. The proposed framework addresses this gap by incorporating psychological and behavioural craving dimensions, thereby providing a more comprehensive representation of stress.

Table 6. Performance comparision of existing and proposed ERFML model

Model	Accuracy	Precision	Recall	F1-score
SVM	0.91	0.90	0.89	0.89
KNN	0.88	0.87	0.86	0.86
Neural Network	0.92	0.91	0.91	0.91
Random Forest	0.94	0.93	0.92	0.92
Proposed ERFML model	0.95	0.94	0.94	0.93

The ROC analysis of the output of various ML classifier models, as presented in the plot figure 3(a), shows that the proposed fusion-based ERFML model achieves the highest AUC (~0.95), indicating superior discriminative capability. While the classification capability of the model is interpreted by the confusion matrix, as shown in Figure 3(b), it reveals a high number of correct classifications (377 out of 397 test samples) with minimal and balanced misclassifications.

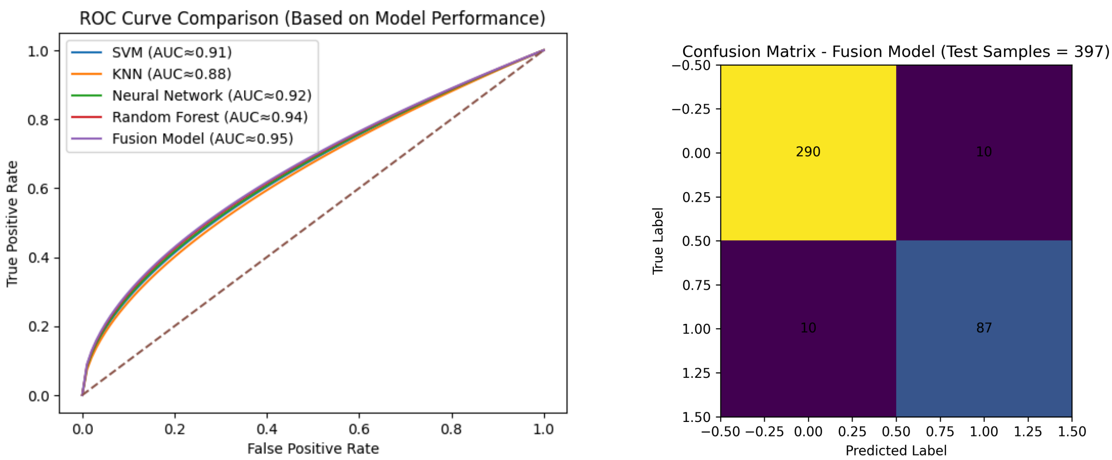

Figure 3. (a) ROC curves comparing the performance of existing models and proposed ERFML model. (b) Confusion matrix of proposed ERFML model

Furthermore, explainability analysis was performed to identify features that contribute most significantly to model predictions. The SHAP-based values, as shown in Figure 4(a), suggest that psychological features of PSS stress remain the most influential factor, followed by physiological features like HR and GSR. These findings are well in line with the fusion-based proposed ERFML feature importance results shown in Figure 4 (b). These insights confirm that both results are aligned and jointly contribute to the model's predictive accuracy and to the effectiveness of multimodal data fusion and integration within the ML framework.

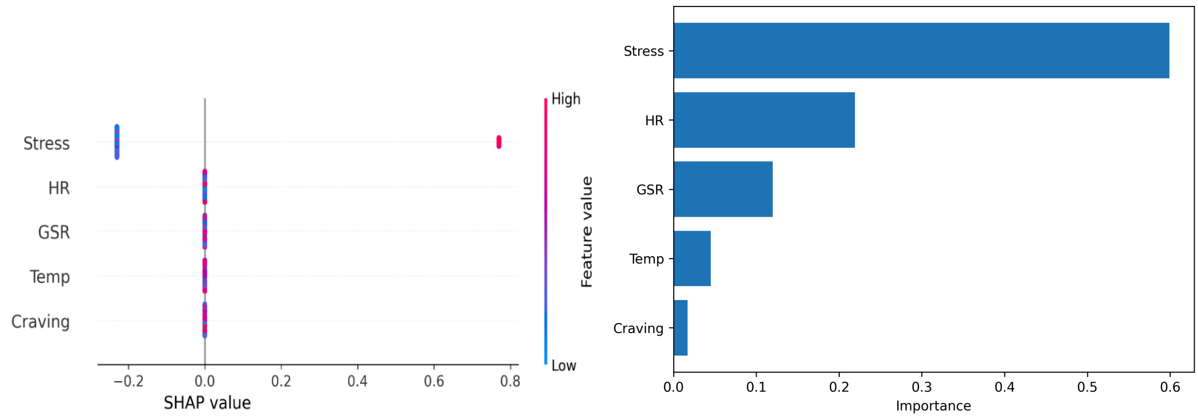

Figure 4. (a) Explainability analysis SHAP output and (b) Feature importance of the of the proposed ERFML model

Furthermore, LOSO cross-validation results demonstrate that the proposed model achieves high performance across unseen participants, with an average accuracy of 91% and low variance. This indicates that the model effectively captures generalisable multimodal patterns of stress and craving rather than memorising individual specific characteristics.

These findings of experimental analysis have evidently found that stress, particularly when acting as a cue for substance consumption, can be more effectively identified through fusion multimodal approach. The proposed ERFML approach not only enhances predictive performance but also addresses a critical gap in stress detection research by linking physiological stress responses with psychological perception and substance craving behaviour, thereby giving insights into the level of attention required to formulate prevention and treatment strategies.

4. Discussion

This study provides sufficient evidence that incorporating a multimodal data acquisition framework for stress monitoring via wearable sensors and psychological assessments can significantly improve detection and intervention for substance use risk due to stress. The proposed study integrates physiological and psychological stress indicators into an early warning system for addiction. It is noted that existing studies often focus on physiological or psychological stress alone, rather than combining the two aspects (Kirsch & Lippard, 2022; Sood et al., 2013), and such a gap in execution on unimodal systems has been overcome by the proposed ERFML model. From the experimental study, it has been observed that fusion of the Multimodal data acquisition framework and the ensemble of ML models remain central to achieving high predictive accuracy in complex behavioural health domains.

Behavioural health and data-driven research have converged over the past decade; making significant progress in identifying, tracking, and predicting psychological factors contributing to substance use and addiction, specific in time-series analysis of psychological datasets (Wang et al., 2016; Bahador et al., 2021). The widespread recognition of stress as a precursor and consequence of drug and alcohol abuse is among these factors. Wearable biosensors, ML, and multimodal data collection are increasingly being used to predict addiction risk and personalise intervention based on the growing understanding of stress as a dynamic and measurable indicator (Rumbut et al., 2022). This two-way interaction emphasises the importance of accurate stress monitoring in detecting individuals at risk for substance abuse or relapse. Individuals recovering from addiction can be tracked using a wearable sensor device in a clinical therapeutic setting. When patients are experiencing withdrawal symptoms or cravings, physiological markers of stress are commonly elevated, including elevated heart rates, aberrant GSRs, and elevated skin temperatures, reflecting psychological stress. Real-time identification of acute stress episodes is necessary to intervene immediately (Carreiro et al., 2024). It is possible to intervene during craving spikes before they lead to substance use. In this case, wearable technology allows for prompt, personalised addiction treatment interventions based on physiological stress expressions. Temperature, measured using the MLX90614 sensor, is included as a physiological feature; however, it has limitations in stress measurement due to slow responsiveness, environmental influence, and variability caused by other health conditions. Although temperature measurement has such limitations, it is still considered part of the model as an additional physiological variable. In the proposed ERFML framework, temperature contributes to the composite stress score but is not relied upon as a dominant predictor. The fusion model assigns relatively lower importance to temperature, as confirmed by feature importance and SHAP analysis.

By providing contextual, real-time feedback about an individual's stress state, the proposed multimodal system bridges the gap between theoretical understanding and clinical application. The model accurately captures and classifies physiological markers, such as increased HR and GSR during cravings or withdrawal. Furthermore, PSS scores are effective in detecting psychological triggers that are often not captured by biosignal analysis alone. Existing research on stress and cues to substance consumption has confirmed that stress is often associated with drug-seeking behaviours (Kulman et al., 2021; Li et al., 2024; Rumbut et al., 2023). The proposed model indicates that more than 70% of detected stress spikes predict craving episodes across various recovery contexts, underscoring the potential for proactive digital intervention. Despite initial challenges related to sensor calibration and data synchronisation, compliance with ethical protocols, including data anonymisation and informed consent, was ensured. These outcomes highlight the promise of multimodal stress detection as a scalable, interpretable, and ethically grounded approach to addiction recovery. Based on the attention level classified by the model, the decision-support system can trigger suggestive treatment plans and interventions.

5. Conclusion

Stress and substance use cues are inherently complementary, and addressing both through early intervention is essential to prevent the disorder from escalating into a severe clinical condition. This study demonstrates that integrating physiological, psychological, and craving-related features through a fusion-based ensemble framework significantly enhances the understanding and prediction of stress-induced substance use behaviour. The proposed ERFML model even effectively overcomes the limitations of traditional single-modality approaches by enabling real-time, context-aware analysis of multimodal data collected from wearable sensors and validated psychological assessments.

The experimental results confirm that stress plays a critical role in influencing substance use patterns. A positive correlation between craving and stress was observed, alongside associations with physiological signals such as heart rate and skin responses. Furthermore, the multimodal approach is strongly recommended, as it outperformed existing models by a substantial margin in predictive performance. With a recall value nearing 95%, it further validates the robustness and reliability of the proposed approach in identifying vulnerability to relapse. Additionally, it has been found to be evident that more than 50% of the withdrawal episodes depend on the craving intensity. This indicates a close association between stress-related craving intensity and relapse. Thus, it makes it necessary to incorporate stress monitoring into addiction treatment frameworks. While the proposed framework demonstrates strong potential for integration, practical deployment would require compliance with healthcare data standards (e.g., interoperability protocols, data security, and privacy regulations).

In this multimodal approach, the dataset acquired is limited in terms of geographical and demographics, however, it performs comparably to the benchmark WESAD dataset. This shows strong implications for the development of intelligent, wearable-assisted intervention systems that can support clinicians in delivering personalised and timely treatment strategies. It provides a scalable and interpretable solution for early detection of high-risk states in an individual's recovery session. Although high accuracy has been obtained, the research has some constraints due to the fact that the results were produced using one dataset under controlled conditions. Hence, it might have some good features that resulted in a good AUC. In future, the model can be tested with varied population and environmental conditions. Furthermore, researchers can expand the dataset to improve generalisability by incorporating longitudinal analysis across broader populations. On top of this, the model can be integrated with adaptive intervention mechanisms such as reinforcement learning to further improve predictive accuracy and real-world applicability.

Acknowledgment

The authors are thankful for the support and resources offered by Centre for Addiction Medicine (CAM), Sikkim, 737126, India. We offer our sincere appreciation to Dr Satish Rasaily, Head of the Department of CAM centre and entire nursing fraternity, students for their data collection efforts and participants for their sincere efforts in making this research fruitful.

References

Al-Atawi, A. A., Alyahyan, S., Alatawi, M. N., Sadad, T., Manzoor, T., Farooq-i-Azam, M., & Khan, Z. H. (2023). *Stress monitoring using machine learning, IoT and wearable sensors*. Sensors, 23(21), 8875. https://doi.org/10.3390/s23218875

Bahador, N., Jokelainen, J., Mustola, S., & Kortelainen, J. (2021). *Multimodal spatio-temporal-spectral fusion for deep learning applications in physiological time series processing: A case study in monitoring the depth of anesthesia*. Information Fusion, 73(1) 125–143. https://doi.org/10.1016/j.inffus.2021.03.001

Balbim, G. M., Marques, I. G., Marquez, D. X., Patel, D., Sharp, L. K., Kitsiou, S., & Nyenhuis, S. M. (2021). *Using Fitbit as an mHealth intervention tool to promote physical activity: Potential challenges and solutions*. JMIR mHealth and uHealth, 9(3), e25289. https://doi.org/10.2196/25289 Bolpagni, M., Pardini, S., Dianti, M., & Gabrielli, S. (2024). *Personalized stress detection using biosignals from wearables: A scoping review*. Sensors, 24(10), 3221. https://doi.org/10.3390/s24103221

Carreiro, S., Ramanand, P., Taylor, M., Leach, R., Stapp, J., Sherestha, S., Smelson, D., & Indic, P. (2024). *Evaluation of a digital tool for detecting stress and craving in SUD recovery: An observational trial of accuracy and engagement*. Drug and Alcohol Dependence, 261, 111353. https://doi.org/10.1016/j.drugalcdep.2024.111353

Carter, T. A., Heaton, K., Merlo, L. J., Roche, B. T., & Puga, F. (2023). *Relapse prevention and prediction strategies in substance use disorder: A scoping review*. Journal of Addictions Nursing, 34(2), 146–157. https://doi.org/10.1097/jan.0000000000000527

Dankovich, L. J., IV, Joyner, J. S., He, W., Sesay, A., & Vaughn-Cooke, M. (2024). CogWatch: An open-source platform to monitor physiological indicators for cognitive workload and stress. *HardwareX, 19*, e00538. https://doi.org/10.1016/j.ohx.2024.e00538

Duffing, T. M., Greiner, S. G., Mathias, C. W., & Dougherty, D. M. (2014). Stress, substance abuse, and addiction. In C. M. Pariante & M. D. Lapiz-Bluhm (Eds.), *Behavioral neurobiology of stress-related disorders* (pp. 237–263). Springer. https://doi.org/10.1007/7854_2014_276

Habelt, B., Arvaneh, M., Bernhardt, N., & Minev, I. (2020). Biomarkers and neuromodulation techniques in substance use disorders. *Bioelectronic medicine*, *6*, 4. https://doi.org/10.1186/s42234-020-0040-0

Jambhale, K., Mahajan, S., Rieland, B., Banerjee, N., Dutt, A., Kadiyala, S. P., & Vinjamuri, R. (2022). *Identifying biomarkers for accurate detection of stress*. Sensors, 22(22), 8703. https://doi.org/10.3390/s22228703

Kargarandehkordi, A., Slade, C., & Washington, P. (2024). *Personalized AI-driven real-time models to predict stress-induced blood pressure spikes using wearable devices: Proposal for a prospective cohort study*. JMIR Research Protocols, 13(1), e55615. https://doi.org/10.2196/55615

Kirsch, D. E., & Lippard, E. T. (2022). *Early life stress and substance use disorders: The critical role of adolescent substance use*. Pharmacology Biochemistry and Behavior, 215, 173360. https://doi.org/10.1016/j.pbb.2022.173360

Koob, G. F., & Volkow, N. D. (2016). *Neurobiology of addiction: A neurocircuitry analysis*. The Lancet Psychiatry, 3(8), 760–773. https://doi.org/10.1016/S2215-0366(16)00104-8

Koob, G. F., Kandel, D. B., Baler, R. D., & Volkow, N. D. (2023). *Neurobiology of addiction*. In *Tasman's psychiatry* (pp. 1–51). Springer International Publishing.

Kulman, E., Chapman, B., Venkatasubramanian, K., & Carreiro, S. (2021). Identifying opioid withdrawal using wearable biosensors. In *Proceedings of the Annual Hawaii International Conference on System Sciences* (Vol. 54, pp. 3583–3592).

Li, S., Fan, C., Kargarandehkordi, A., Sun, Y., Slade, C., Jaiswal, A., Benzo, R. M., Phillips, K. T., & Washington, P. (2024). *Monitoring substance use with Fitbit biosignals: A case study on training deep learning models using ecological momentary assessments and passive sensing*. AI, 5(4), 2725–2738. https://doi.org/10.3390/ai5040131

Mimura, C., & Griffiths, P. (2003). *The effectiveness of current approaches to workplace stress management in the nursing profession: An evidence-based literature review*. Occupational and Environmental Medicine, 60(1), 10–15. https://doi.org/10.1136/oem.60.1.10

Neigel, P., Vargo, A., Tag, B., & Kise, K. (2024, October). *Using wearables to unobtrusively identify periods of stress in a real university environment*. In *Proceedings of the 2024 ACM International Symposium on Wearable Computers* (pp. 17–24). ACM. https://doi.org/10.1145/3675095.3676608

Nikbakhtzadeh, M., Ranjbar, H., Moradbeygi, K., Zahedi, E., Bayat, M., Soti, M., & Shabani, M. (2023). *Cross-talk between the HPA axis and addiction-related regions in stressful situations*. Heliyon, 9(4), e15525. https://doi.org/10.1016/j.heliyon.2023.e15525

Ovsiannikova, Y., Pokhilko, D., Kerdyvar, V., Krasnokutsky, M., & Kosolapov, O. (2024). *Peculiarities of the impact of stress on physical and psychological health*. Multidisciplinary Science Journal, 6. https://doi.org/10.31893/multiscience.2024ss0711

Razavi, M., McDonald, A., Mehta, R., & Sasangohar, F. (2023). *Evaluating mental stress among college students using heart rate and hand acceleration data collected from wearable sensors*. arXiv. https://arxiv.org/abs/2309.11097

Ruisoto, P., & Contador, I. (2019). *The role of stress in drug addiction: An integrative review*. Physiology & Behavior, 202, 62–68. https://doi.org/10.1016/j.physbeh.2019.01.022

Rumbut, J., Fang, H., Carreiro, S., Smelson, D., & Boyer, E. (2022). *An overview of wearable biosensor systems for real-time substance use detection*. IEEE Internet of Things Journal, 9(23), 23405–23415. https://doi.org/10.1109/JIOT.2022.3207090

Rumbut, J., Fang, H., Boyer, E. W., & Wang, H. (2023, June). *Short: Precision polysubstance use episode detection in wearable biosensor data streams*. In *Proceedings of the 8th ACM/IEEE International Conference on Connected Health: Applications, Systems and Engineering Technologies* (pp. 163–167). ACM. https://doi.org/10.1145/3580252.3586989

Schmidt, P., Reiss, A., Duerichen, R., Marberger, C., & Van Laerhoven, K. (2018, October). *Introducing WESAD, a multimodal dataset for wearable stress and affect detection*. In *Proceedings of the 20th ACM International Conference on Multimodal Interaction* (pp. 400–408). ACM. https://doi.org/10.1145/3242969.3242985

Schwabe, L., Dickinson, A., & Wolf, O. T. (2011). *Stress, habits, and drug addiction: A psychoneuroendocrinological perspective*. Experimental and Clinical Psychopharmacology, 19(1), 53–63. https://doi.org/10.1037/a0022212

Shrestha, S., Stapp, J., Taylor, M., Leach, R., Carreiro, S., & Indic, P. (2023). Towards device agnostic detection of stress and craving in patients with substance use disorder. In *Proceedings of the Annual Hawaii International Conference on System Sciences* (pp. 3156–3163). https://doi.org/10.24251/HICSS.2023.388

Sinha, R. (2008). *Chronic stress, drug use, and vulnerability to addiction*. Annals of the New York Academy of Sciences, 1141(1), 105–130. https://doi.org/10.1196/annals.1441.030

Sinha, R. (2024). Stress and substance use disorders: Risk, relapse, and treatment outcomes. *The Journal of Clinical Investigation, 134*(16), e172883. https://doi.org/10.1172/JCI172883

Sun, Y., Kargarandehkordi, A., Slade, C., Jaiswal, A., Busch, G., Guerrero, A., Phillips, K. T., & Washington, P. (2024). Personalized deep learning for substance use in Hawaii: Protocol for a passive sensing and ecological momentary assessment study. *JMIR Research Protocols, 13*(1), e46493. https://doi.org/10.2196/46493

Sood, P., Priyadarshini, S., & Aich, P. (2013). *Estimation of psychological stress in humans: A combination of theory and practice*. PLOS ONE, 8(5), e63044. https://doi.org/10.1371/journal.pone.0063044

Takano, A., Ono, K., Nozawa, K., Sato, M., Onuki, M., Sese, J., Yumoto, Y., Matsushita, S., & Matsumoto, T. (2023). *Wearable sensor and mobile app-based mHealth approach for investigating substance use and related factors in daily life: Protocol for an ecological momentary assessment study*. JMIR Research Protocols, 12, e44275. https://doi.org/10.2196/44275

Tavolacci, M. P., Ladner, J., Grigioni, S., Richard, L., Villet, H., & Dechelotte, P. (2013). *Prevalence and association of perceived stress, substance use and behavioral addictions: A cross-sectional study among university students in France, 2009–2011*. BMC Public Health, 13(1), 724. https://doi.org/10.1186/1471-2458-13-724

Vavrinsky, E., Stopjakova, V., Kopani, M., & Kosnacova, H. (2021). *The concept of advanced multi-sensor monitoring of human stress*. Sensors, 21(10), 3499. https://doi.org/10.3390/s21103499

Wang, K., Zhao, Y., Xiong, Q., Fan, M., Sun, G., Ma, L., & Liu, T. (2016). *Research on healthy anomaly detection model based on deep learning from multiple time-series physiological signals*. Scientific Programming, 2016, 5642856. https://doi.org/10.1155/2016/5642856

Xefteris, V.-R., Dominguez, M., Grivolla, J., Tsanousa, A., Zaffanela, F., Monego, M., Symeonidis, S., Diplaris, S., Wanner, L., Vrochidis, S., & Kompatsiaris, I. (2023). *A multimodal late fusion framework for physiological sensor and audio-signal-based stress detection: An experimental study and public dataset*. Electronics, 12(23), 4871. https://doi.org/10.3390/electronics12234871