Nottingham Trent University

Department of Computer Science

Spatial and Temporal Environment Impact Analysis on People's Well-being

Faiza Guerrache

School of Science & Technology

A thesis submitted in partial fulfilment of the requirements

for the degree of

Doctor of Philosophy

February 2024

This thesis is dedicated to my beloved parents, whose unwavering support and guidance have been the cornerstone of my academic journey. Your encouragement and invaluable life lessons have paved the way for this achievement, and I am forever grateful for your boundless love and belief in me.

Acknowledgements

First and foremost, I express my deepest gratitude to Allah, the most merciful, for His endless blessings and guidance throughout my postgraduate studies.

I extend my heartfelt appreciation to my supervisory team, Prof. David Brown and Dr Mufti Mahmud, for their unwavering support, invaluable guidance, and encouragement. Their wealth of knowledge and insightful advice have been instrumental in shaping this thesis. I am profoundly grateful for their constant support and patience, which have been pivotal in navigating the challenges of doctoral research.

I am also indebted to Prof. Eiman Kanjo, Dr. Andreas Oikonomou, and Dr. Golnaz Shahtahmassebi for their encouragement and support throughout this journey. Their expertise and encouragement have been invaluable assets.

To my parents, I owe an immeasurable debt of gratitude for their unconditional love, unwavering care, and unwavering support throughout my life. Their encouragement has been a constant source of strength and inspiration.

To my dear children, I am profoundly grateful for your unwavering support, encouragement, and patience throughout this doctoral endeavour. Your presence has been a constant source of motivation and joy.

I extend my sincere thanks to all those who have supported and encouraged me along this journey. Your contributions, whether big or small, have been deeply appreciated and will always be remembered.

Abstract

This PhD research programme presents an innovative approach to understand the environmental factors of human wellbeing through the development and analysis of the "EnviroWellBeing" dataset and the application of advanced machine learning and deep learning algorithms for environment and stress-level classification tasks. The careful curation of the dataset involved the synchronisation of sensor data to a uniform 1Hz frequency and the application of comprehensive data cleaning processes, ensuring its suitability for time-series analysis. The dataset curation effort marks a significant advancement in studying the spatial and temporal impacts of environmental factors on physiological and psychological states.

Additionally, this research explores the "Depresjon" dataset, applying data analysis techniques to uncover patterns in motor activity related to depression. A comparative analysis of machine learning models demonstrated the ability to distinguish between depressed and healthy individuals using motor activity, with a Random Forest (RF) classifier achieving 83.41% accuracy. Analysis of the Depresjon dataset reveals key physiological markers of depression and highlights the role of predictive modeling in advancing mental health research.

Using the EnviroWellBeing Dataset, the research details the higher performance of 1D Convolutional Neural Networks (1D-CNNs), which achieved notable accuracies in classifying environmental conditions (e.g., 97.72% in situ, 94.18% in vitro) and stress levels (e.g., 82.37% in situ, 63.37% in vitro), highlighting their effectiveness. The evaluation also includes the effectiveness of Long Short-Term Memory (LSTM) networks in capturing sequential dependencies and the robustness of RF classifiers as a non-sequential baseline. Key findings demonstrate the critical role of feature selection (identifying CO2, wrist temperature, and NO as key predictors in situ), the models' capacity to generalise from physiological responses to stress, and the provision of valuable insights into feature importance for future model development.

By offering valuable insights into the performance of machine learning models in environmental and stress-level classifications, alongside a comprehensive dataset, this PhD research programme significantly contributes to the fields of environmental health and mental wellbeing. The findings demonstrate potential applications in urban planning and personal health monitoring by showing how sensor data analysis can inform strategies to mitigate environmental stressors and enhance human health and happiness.

Publications

The following publications are a result of the research presented in this thesis.

1. Guerrache, F., David, J.B., and Mufti M. Urban-Health: The relationship between environment factors and heart health. *submitted to Noise and Health Journal (2024).*

2. Guerrache, F., David, J.B., and Mufti M. The "EnviroWellBeing" dataset as a publicly accessible dataset. *submitted to Data-in-Brief (2024).*

3. Guerrache, F., David, J.B., and Mufti M. Classifying Depressed and Healthy Individuals Using Wearable Sensor Data: A Comparative Analysis of Machine Learning and Deep Learning Approaches. *Accepted for publication by the 25th International AIDS Conference (2024).*

4. Guerrache, F., David, J.B., and Mufti M. Impacts of Environmental Factors on Wellbeing: Machine Learning based Benchmarking of Spatial and Temporal Properties. *Accepted for publication by ICICT (2024).*

5. Guerrache, F., David, J.B. and Mufti M. Impacts of Environmental Factors on Wellbeing. Machine Learning based Benchmarking of Spatial and Temporal Properties. *In ICICT (2024).*

6. Guerrache, F., David, J.B., and Mufti M. Classifying Depressed and Healthy Individuals Using Wearable Sensor Data: A Comparative Analysis of Machine Learning Approaches. *In 3rd International Conference on Applied Intelligence and Informatics (AII) (2023).*

7. Guerrache, F., Aldabbagh, A., and Kanjo, E. Multiple sensor fusion approach to map environmental noise impact on health. *In Proceedings of the ACM International Joint Conference on Pervasive and Ubiquitous Computing. Adjunct (2016) 1074-1078.*

8. Guerrache, F., Younis, E.M., and Kanjo, E. Quantifying Environmental Noise Impact on Heart Rate Variability. *In MobiCASE (2016) 149-150.*

Nomenclature

Acronyms

1D-CNN	One-Dimensional Convolutional Neural Network used for extracting patterns from sequential or time-series data.
Autoencoder	A neural network that learns compressed representations of data, often used for dimensionality reduction or noise filtering.
CNN	Convolutional Neural Network; a deep learning model applying convolutional filters to grid-like data.
DL	DL Deep Learning; a subset of machine learning that uses neural networks with many layers to model complex patterns in data.
DTW	Dynamic Time Warping; algorithm to measure similarity between time-series that may vary in speed.
ECG	Electrocardiogram; records electrical activity of the heart, used to assess heart health and stress response.
EDA	Electrodermal Activity; measures skin conductance linked to emotional and physiological arousal.
Encoder	Part of an autoencoder that transforms input data into a compressed (latent) representation.

EnviroWellBeing Dataset	A dataset developed in this thesis containing synchronised environmental and physiological data from multiple sensor sources and participants.
FCN	Fully Convolutional Network; a neural network consisting entirely of convolutional layers, used for spatial predictions.
GRU	Gated Recurrent Units: A simplified type of recurrent neural network architecture that uses gating mechanisms to control the flow of information, often used for time-series and sequential data analysis.
GB	Gradient Boosting: an ensemble machine learning technique that builds models sequentially to correct the errors made by previous models.
GIS	Geographic Information Systems: tools and frameworks for handling and analysing geographic and spatial data.
HR	Heart Rate; number of heartbeats per minute, an indicator of physiological state.
KNN	K-Nearest Neighbours a classification method that assigns labels based on the majority label among the k closest data points.
LR	Logistic Regression; a statistical model used for binary classification that predicts the probability of a categorical dependent variable.
LSTM	Long Short-Term Memory; a type of recurrent neural network effective for learning long-term dependencies in sequential data.

ML	Machine Learning; a field of artificial intelligence that enables systems to learn patterns from data and make predictions or decisions without being explicitly programmed.
MLP	Multi-Layer Perceptron; a type of artificial neural network with multiple layers used for supervised learning.
MLSTM-FCN	A hybrid model combining Multivariate LSTM and Fully Convolutional Network for time-series classification tasks.
PCA	Principal Component Analysis; a dimensionality reduction technique that transforms data into a set of uncorrelated components.
RF	Random Forest; an ensemble machine learning algorithm that constructs multiple decision trees and merges their results to improve classification and regression performance.
RNN	RNN Recurrent Neural Network; a class of neural networks where connections between units form directed cycles, allowing temporal dynamic behaviour suitable for sequential data.
ROC Curve	Receiver Operating Characteristic Curve; shows the performance of a classification model at all classification thresholds.
Sensor Fusion	The integration of data from multiple sensors to provide more accurate or comprehensive information.
Stress Level Classification	A supervised learning task to determine the stress level (e.g., relaxed, neutral, stressed) from sensor data.

Time Series Classification	Assigning class labels to sequential data over time using machine or deep learning techniques.
TPOT	Tree-based Pipeline Optimisation Tool; an automated machine learning (AutoML) tool that optimises machine learning pipelines using genetic programming.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

1.1 Motivation and Research Background

Understanding the influence of environmental factors on human wellbeing is crucial, as physical and mental health are closely linked to the spatial and temporal aspects of our surroundings. Although many studies have examined variables like air quality or green space, these typically focus on isolated factors rather than exploring how multiple environmental dimensions interact over time and space. This fragmented approach limits a full understanding of how complex human–environment interactions affect wellbeing.

Advances in sensor technology and analytical methods, such as deep learning, now enable the collection and integration of multimodal data—including environmental, biometric, and contextual information—in real time. Leveraging these technologies offers a unique opportunity to study the combined and dynamic effects of environmental exposures on physiological and psychological health. This PhD research programme aims to fill this gap by developing a comprehensive approach that synthesises spatial and temporal environmental data with human physiological measures, thereby providing clearer insights into how different environments shape wellbeing.

Addressing this gap is essential for improving urban planning and public health interventions aimed at promoting healthier living environments.

1.2 Aim and Objectives

This PhD research programme aims to develop a comprehensive understanding of how environmental conditions, in both urban and green spaces, affect human wellbeing. The research tests the hypothesis that the spatial and temporal dimensions of an environment significantly shape physiological and psychological health. The study seeks to provide significant insights into the interplay between surroundings and wellbeing state.

Achieving this aim requires completing a set of SMART objectives:

- To review and synthesise existing literature on multimodal and data fusion techniques relevant to time-series sensor data analysis.
- To develop and validate a multi-sensor data fusion architecture capable of integrating environmental, biometric, and contextual data.
- To design and conduct at least three experimental studies across both indoor (in vitro) and outdoor (in situ) environments, capturing physiological responses to a minimum of four distinct environmental stimuli.
- To apply advanced data preprocessing techniques to ensure data quality and readiness for analysis throughout the data collection phase.
- To implement and evaluate machine learning models that identify correlations between environmental conditions and wellbeing indicators, achieving at least 80% classification accuracy.
- To develop data visualisation tools that effectively represent the impact of environmental factors on wellbeing, completing prototypes.

These objectives guide this comprehensive investigation. The research aims to address significant gaps in understanding how the surroundings influence health and happiness. Moreover, this work endeavours to foster multidisciplinary dialogue among urban planners, public health officials, and environmental scientists. The goal is to extend the impact beyond the academic arena, informing urban development and public health strategies. This ambition underlines the

potential to promote environments that support sustainable and healthy lifestyles, illustrating the importance of integrating environmental factors into the broader discourse on human wellbeing.

1.3 Research Questions

This research is driven by several key questions, each designed to probe the intricate relationship between environmental factors and human wellbeing:

RQ1: Can a practical method be developed to collect and integrate multimodal sensor data from environmental, biometric, and contextual sources across different experimental settings, including polluted and green outdoor spaces (in situ) and controlled indoor environments with varying sound conditions (in vitro), to facilitate a comprehensive analysis of environmental impact on wellbeing?

RQ2: Can physiological data alone accurately differentiate between environmental conditions, such as polluted versus green outdoor spaces and exposure to annoying versus pleasant sounds in controlled settings?

RQ3: How effectively can integrated environmental and physiological data predict individuals' stress levels, and what role do specific environmental factors and physiological measures play in enhancing the accuracy of these predictions?

RQ4: Is time-series data derived from environmental and physiological sensors more insightful compared to current, snapshot information in predicting stress levels, and what implications does this have for the temporal dynamics of stress response?

1.4 Contributions

This PhD research programme advances the field of environmental health and wellbeing analytics by developing an integrated approach to the collection, analysis, and interpretation of multimodal sensor data. The research begins with a comparative analysis of public mental health data in Chapter 3, applying classical machine learning algorithms to the Depresjon dataset to objectively classify depressive states using wearable sensor signals. This early investigation

establishes methodological foundations for feature selection and classification, as well as highlighting practical considerations relevant to sensor-based wellbeing studies.

Building on these foundations, the thesis introduces the EnviroWellBeing dataset in Chapter 4—a comprehensive and novel resource that combines synchronized physiological, environmental, and contextual data captured in both real-world and controlled experimental settings. The scale and granularity of this dataset support in-depth analyses of environmental exposures and human physiological responses over time and space.

Methodologically, the research contributes robust preprocessing pipelines and segmentation strategies that harmonize heterogeneous time-series data from multiple sensors. It systematically implements and benchmarks a suite of machine learning and deep learning frameworks in Chapters 5 and 6, including temporal models such as LSTM and 1D-CNNs, for both environmental classification and stress state prediction. These approaches enable the integration of multimodal signals, addressing challenges that have limited previous sensor-based research.

Empirical analyses identify critical physiological and environmental predictors of stress, using systematic feature importance techniques to interpret model outputs and guide future research directions. The thesis also demonstrates in Chapter 6 the feasibility of practical, wearable-only stress monitoring systems, laying a foundation for applications in personal health management and policy development. Taken together, these contributions establish a rigorous, data-driven framework for advancing the understanding of environmental influences on human wellbeing.

1.5 Thesis Outline

This thesis is organised into seven chapters following this Introduction (Chapter 1). Chapter 2 merges the state-of-the-art review with the theoretical background, covering environmental impacts on wellbeing, the role of technology, and the foundational machine learning and deep learning concepts used in the research. Chapter 3 details the methodology and findings from classifying mental health states using the public dataset. Chapter 4 describes the creation, preprocessing, and analysis of the bespoke dataset developed for this research. Chapter 5 presents the application of machine learning and deep learning models to the dataset for environmental

and stress-level classification. Chapter 6 discusses the development and results of a framework for predicting physiological states using sensor data. Finally, Chapter 7 summarises the key findings and contributions of the thesis, discusses implications, addresses limitations, and suggests directions for future research.

Chapter 2

Background and State-of-the-Art

2.1 Introduction

This chapter provides a comprehensive background and review of the state-of-the-art relevant to understanding the impact of environmental factors on human wellbeing. It begins by outlining the significance of environmental influences and introduces key theoretical concepts from environmental psychology and spatial-temporal data analysis. The chapter then discusses the crucial role enabling technologies, particularly sensors and data analysis methods, play in this research area. Following this, it delves into specific machine learning and deep learning approaches applicable to environmental wellbeing analysis, covering algorithms, data considerations, and feature engineering. The chapter also presents applications and case studies from the literature, explores pertinent challenges including ethical considerations, and concludes by summarising the identified research gaps that this PhD research programme aims to address.

2.2 Environmental Impacts on Wellbeing

The wellbeing of individuals and communities is profoundly shaped by environmental factors possessing both spatial and temporal dimensions. Recent studies analyse how dynamic changes in ecological, urban, climate, and biodiversity indicators across locations and over time impact human experiences and social outcomes [1]. Urban design elements including den-

sity, green spaces vegetation, and commerce significantly influence wellbeing. These elements affect stress, restoration, and social integration. Meanwhile, factors like industrial structure, financial development, and urbanisation drive ecological welfare performance across cities and regions in complex ways [2]. The spatial distribution of biodiversity also changes over time, with consequences for ecosystem health and human wellbeing [3]. Furthermore, the temporal evolution of ecological and environmental factors creates dynamic changes in wellbeing over years and decades. Variables like rainfall, temperature, and sunshine demonstrate temporal impacts on life satisfaction [4]. The complex interplay between space, time, and environmental factors also mediates outcomes such as property crime rates, underscoring indirect social influences [5]. Recent evidence clearly highlights that human wellbeing is contingent on the quality of, and changes in, the surrounding environmental context across both spatial and temporal dimensions. As such, incorporating a spatio-temporal understanding of environmental factors is critical. This understanding informs urban design, regional development, and biodiversity conservation policies aimed at promoting sustainable wellbeing.

Analysing these environmental impacts is a cornerstone for public health advancement and the sustainable development of urban landscapes. Research findings underscore the significance of integrating environmental impact assessments within public health and urban planning, aiming to foster sustainable, resilient, and health-oriented communities [6–8]. The intricate relationship between environmental changes and public health outcomes necessitates a comprehensive, interdisciplinary approach. Lauriola et al. [6] emphasise the critical role of Environmental and Public Health Tracking (EPHT) systems. These systems amalgamate data across domains to clarify connections between environmental hazards, exposures, and health outcomes, enabling precise, efficient, and cost-effective public health interventions. The global EPHT network exemplifies the potential for international collaboration to enhance understanding and management of environmental health risks, aligning with sustainable development goals.

Environmental exposures are complex, often involving low doses of multiple chemical pollutants, demanding integrated approaches beyond simple exposure-disease paradigms. Arebola et al. [7] advocate for exploring the full exposure-disease continuum using insights from epidemiology, basic sciences, and clinical research. This integrated methodology is pivotal for

unravelling nuanced interactions between pollutants and health, informing public health campaigns. The socio-economic dimensions are also critical; novel methodologies like cost-benefit analyses are needed to grasp broader impacts and craft policies addressing environmental health disparities.

Turning to urban planning, integrating environmental impact assessments is fundamental. Sustainable Urban Mobility Plans (SUMPs) can mitigate transport externalities [8], while Life Cycle Assessment (LCA) methodologies help align development with sustainability objectives [9]. Urban planning decisions significantly influence residents' wellbeing [10]. Strategies enhancing social services, community engagement, and green spaces improve urban livability [10]. Furthermore, methods like the Hedonic Pricing Method (HPM) show the economic value residents place on open spaces, reinforcing the need to prioritise them [11].

The confluence of environmental health and urban planning research underscores the critical importance of analysing environmental impacts. Interdisciplinary approaches, integrated epidemiology, and attention to socio-economic determinants are paramount. Concurrently, incorporating environmental considerations into urban planning is indispensable for sustainable urban growth, enhanced quality of life, and ecological resilience. This synthesis reflects a convergence of insights from diverse research endeavours. It also advocates for a holistic and integrated perspective when tackling the environmental determinants of health and urban wellbeing.

2.3 Theoretical Foundations

2.3.1 Environmental Psychology

A robust multidisciplinary evidence base illuminates the intricate interplay between the environment and human health, behaviour, and wellbeing outcomes [12–17]. Fundamental theories in environmental psychology offer initial frameworks for conceptualising this complex dynamic. The prevalent Biophilia Hypothesis postulates an innate human biological inclination towards nature, rooted in evolutionary processes. Empirical observations, such as improved af-

fect and cognition following exposure to natural stimuli, support the idea that this affinity arose from fitness advantages associated with monitoring the natural landscape. Contemporary extensions like Nature Connectedness examine the emotional benefits derived from feeling close to the natural world. Quantifications of nature connectedness correlate with mental wellbeing measurements, suggesting that interdependence with nature contributes to flourishing.

Beyond enhancing positive dimensions, natural environments also regulate neural networks involved in emotion processing [14]. Neuroimaging shows that natural scenery properties activate attention, emotion regulation, and default mode brain networks. This neurophysiological cascade subsequently reduces stress and rumination while restoring cognitive capacities. Besides green spaces, elements of built environments like space and light also modulate emotional and cognitive functioning. Virtual environment studies indicate interior spatial geometry impacts vagal regulation of cardiac rhythms via the central autonomic network. Since vagal tone proxies neurocognitive flexibility and emotional wellbeing, environmental design could potentially restore depleted cognitive reserves.

In contrast, environmental stressors like air and noise pollution, high population density, and monotonous architecture have negative effects on wellbeing . Analyses consistently link poor environmental quality exposure with perceived stress, anxiety, depression, cognitive dysfunction, and even more severe psychopathologies [15]. The doubling of mood and anxiety disorder prevalence between 1990 and 2017 coincided with rising global urbanisation; while correlation is not causation, these parallels implicate environmental factors in emotional distress.

Intriguing microbiological mechanisms might also mediate responses. The Lovebug Effect theorises that beneficial microbes from environmental exposures augment stress resilience via the gut-brain axis . Given that gut microbiome richness correlates positively with health and longevity, nurturing a flourishing microbiome through lifestyle and spaces may sustain human flourishing [12].

Finally, the COVID-19 pandemic profoundly increased the influence of immediate home and neighbourhood environments on daily life [17]. Lockdowns channelled activities into domestic spaces, making residential conditions pivotal in mitigating or exacerbating pandemic stres-

sors like financial strain and social isolation. Residential elements like nature views and calming interior designs were found to reduce pandemic-associated depression and anxiety. More broadly, sustainable behaviours and environmental stewardship promote subjective wellbeing through value-action congruency [16]. Choosing spaces and practices that recognise human-environment interdependence can engender coherence and flourishing. Extensive evidence underscores the urgency of integrating insights from architecture, ecology, neuroscience, and psychology to inform lifestyle and policy. Supporting biophilic inclinations through green architecture and conservation can enrich mental health and resilience. Since planetary health depends on human behaviour, propagating spaces and worldviews that facilitate awareness of the person-planet connection is imperative.

2.3.2 Spatial and Temporal Data Analysis

Geographic Information Systems (GIS) are frameworks designed to capture, store, manipulate, analyse, manage, and present spatial or geographic data [18]. GIS integrates hardware, software, and data for handling geospatial information, encompassing applications, data, hardware, and the human element including user expertise. GIS is underpinned by Geographic Information Science principles, which develop data organisation models and computational mechanisms. These include visualisation and spatial analysis techniques essential for understanding spatial relationships. Spatial data models feature the vector model (points, lines, polygons) and the raster model (grid of cells). Spatial analysis techniques involve operations like buffer analysis, overlay analysis, and network analysis, mathematically represented as functions on these models. For example, a buffer analysis creating a zone around point p with distance d can be represented as:

$$B(p, d) = \{q \in S \mid \text{distance}(p, q) \leq d\} \tag{2.1}$$

where S is the space and $distance(p, q)$ is Euclidean distance. Overlay analysis combines data layers; for instance, the intersection of two polygon layers A and B is:

$$A \cap B = \{x \mid x \in A \text{ and } \ x \in B\} \tag{2.2}$$

These representations form the basis for GIS software algorithms.

Time series analysis uses statistical techniques to analyse time-ordered data points, extracting meaningful statistics and characteristics [19]. It is widely used in fields like economics and environmental science to forecast future values based on past observations. The theoretical foundation involves models describing time series behaviour. A fundamental model is the Autoregressive Integrated Moving Average (ARIMA), used to understand temporal dependencies and forecast future values based on the series' own past values, past errors, and degree of differencing. The ARIMA(p, d, q) model is represented as:

$$(1 - \sum_{i=1}^{p} \phi_i L^i)(1 - L)^d X_t = (1 + \sum_{i=1}^{q} \theta_i L^i)\varepsilon_t \tag{2.3}$$

where:

- X_t is the time series value at time t.
- L is the lag operator ($L^k X_t = X_{t-k}$).
- ϕ_i are the parameters of the autoregressive (AR) part (order p).
- θ_i are the parameters of the moving average (MA) part (order q).
- d is the degree of differencing required to make the series stationary.
- ε_t is white noise, typically assumed to be independent and identically distributed from a normal distribution.

This model helps understand temporal dependencies crucial for forecasting.

Integrating GIS with time series analysis allows examining how spatial phenomena change over time. This integration is useful in environmental studies, urban planning, and disaster management, where understanding temporal dynamics of spatial phenomena is crucial. For instance, analysing disease spread involves mapping spatial distribution with GIS and modelling temporal trends with time series analysis to predict future outbreaks.

In conclusion, GIS and time series analysis are powerful tools. Their theoretical foundations, including mathematical representations for spatial data models, analysis techniques, and time

series models, provide a robust framework for analysing complex spatio-temporal phenomena.

2.4 Role of Technology in Environmental Impact Analysis

The integration of technology has significantly transformed the methods used for assessing and mitigating environmental impacts [20, 21]. Moving beyond traditional manual data collection, which was often time-consuming and limited in scope, modern approaches leverage technological advancements for more precise and dynamic analyses crucial for environmental conservation [22, 23].

Information and communication technologies (ICT), advanced computational models, sensor networks, and data analytics have provided new capabilities. These tools allow for the handling of complex, large-scale datasets and support the prediction of environmental impacts with greater reliability compared to older methods [24, 25]. For instance, sensor technologies facilitate real-time monitoring of environmental variables, while computational techniques enable sophisticated modelling and simulation [26]. This technological evolution allows researchers and policymakers to gain deeper insights into environmental processes and the effects of human activities. The effective use of these technologies requires continued innovation and supportive policy frameworks to fully realise their potential for sustainable environmental management. Specific machine learning and deep learning techniques applicable to this research are discussed in Section 2.5.

2.5 Machine Learning and Deep Learning Approaches for Environmental Wellbeing Analysis

2.5.1 Overview of AI, Machine Learning, and Deep Learning

Artificial Intelligence (AI) enables systems to perform tasks typically requiring human intelligence. Machine Learning (ML), a subset of AI, focuses on algorithms that allow systems to learn from data without being explicitly programmed [27, 28]. These algorithms improve

automatically through experience with data. ML is broadly categorised into supervised learning (using labelled data to map inputs to known outputs, including regression and classification tasks relevant to this research), unsupervised learning (finding patterns in unlabelled data, like clustering), and reinforcement learning (learning through rewards and penalties, less directly applicable here) [29].

Deep Learning (DL) is a specialised area within ML that utilises Artificial Neural Networks (ANNs) with multiple layers (deep architectures) to model complex patterns and relationships in large datasets [30]. Unlike traditional ML approaches that might operate linearly or require manual feature engineering, DL models can learn hierarchical features directly from data in a non-linear manner [30]. This capability makes DL particularly powerful for tasks involving complex data such as time-series sensor readings, image analysis, or natural language processing, which are pertinent to analysing environmental impacts on wellbeing. This section will explore specific ML and DL algorithms relevant to analysing the spatio-temporal environmental and physiological data gathered in this research programme.

2.5.2 Data Acquisition and Preprocessing Considerations

Data Acquisition. The proliferation of smartphone applications and wearable sensor devices has unlocked unprecedented opportunities. These technologies allow monitoring environmental exposures and quantifying their impacts on human physiology in real time. Multiple pilot studies have utilised consumer technologies to simultaneously capture personal data, for example using apps like NoiseSpy for noise exposure monitoring [31]. Alongside environmental data, system designs often integrate multiple wearable sensors to track health conditions, capturing physiological indicators like heart rate variability using biometric trackers (See Figure 2.1) [32, 33].

For instance, a field study across Nottingham, UK, equipped participants with portable noise meters and wrist-worn heart rate monitors. This allowed researchers to characterise the immediate effects of urban soundscapes on heart rate patterns and map noise pollution levels to cardiovascular strain, demonstrating the utility of personal sensor data for evaluating population-level

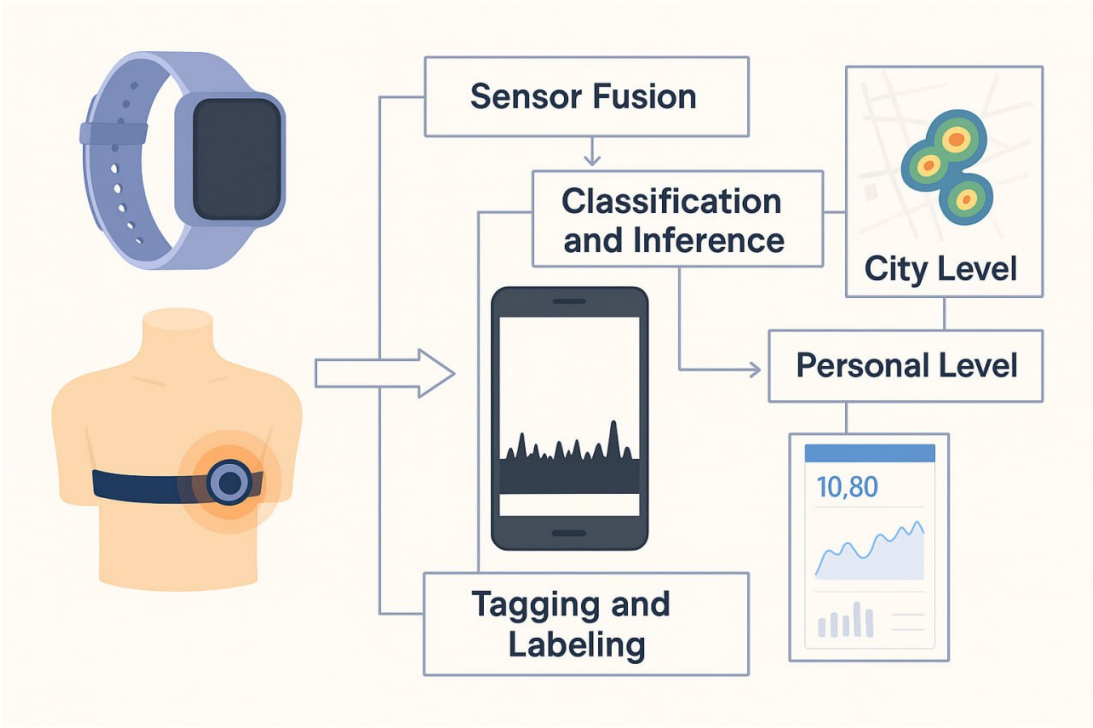


Figure 2.1: System design for tracking health conditions utilising multiple wearable sensors.

environmental health hazards.

Beyond noise, machine learning techniques show increasing promise in forecasting various environmental risk factors linked to health outcomes [34]. Sensor networks now generate vast real-time data on hyperlocal air quality, weather, and other urban health parameters. Integrating this data with satellite imaging, medical records, and crowdsourced reports provides fertile ground for training ML algorithms (including deep neural networks) to model multi-dimensional exposure risks with spatio-temporal precision. For example, hybrid AI architectures combining LSTMs with autoencoders have reliably predicted daily PM2.5 levels from meteorological data, outperforming traditional statistical approaches [35]. Similarly, neural networks have accurately ranked industrial chemicals by anticipated life cycle health impacts based on existing datasets [36]. These examples highlight ML's versatility in forecasting hazards and guiding preventative design decisions.

Overall, integrating ML analysis with diverse environmental monitoring sources provides new tools for mapping and predicting community health threats. However, substantial barriers remain, including fragmented data sources, patient privacy concerns, and the lack of large labelled

training datasets, which currently constrain model accuracy and real-world validity. Addressing these challenges through collaborative data sharing while preserving confidentiality is crucial for advancing practical environmental health forecasting. The potential remains significant for digital health platforms that synthesise environmental tracking, geospatial modelling, and clinical data to proactively defend population health [26].

Data Preprocessing. Recent research highlights crucial advancements in data preprocessing techniques for enhancing ML in environmental health studies. Sun et al. [36] emphasised that data quality and structure significantly impact model performance and interpretability. Advanced preprocessing strategies help address common issues in environmental datasets, such as high dimensionality, missing values, class imbalance, and noisy raw data.

Feature selection is vital for refining environmental data and improving model generalisability. Sun et al. [36] employed a mutual information-permutation importance approach to select relevant molecular descriptors, drastically reducing dimensionality while retaining interpretability and improving accuracy in predicting chemical impacts. Similarly, Dharmasaputro et al. [37] demonstrated that effectively handling missing and imbalanced data using techniques like Multiple Imputation by Chained Equations (MICE) and Synthetic Minority Oversampling Technique (SMOTE) enhanced classifier performance.

In addition to statistical cleaning, deep learning offers new paradigms for automated preprocessing. Kazemzadeh et al. [38] devised cascaded convolutional neural networks capable of inherently correcting defects in Raman spectral data, removing the need for manual baseline correction. Such end-to-end pipelines can improve efficiency and reduce bias. Expanding on this, Bilal et al. [39] proposed Auto-Prep, an interactive architecture that automatically evaluates and recommends optimal preprocessing sequences based on the data. Applied to diverse datasets, this automation simplified workflows and boosted model accuracy.

Innovation in data preprocessing methodologies opens new avenues for advancing ML in environmental health research. Strategies like feature selection, class re-balancing, automated defect correction, and pipeline recommendation cater well to the complexity of environmental data. As these techniques mature, researchers can leverage them with state-of-the-art ML to

uncover deeper environmental health insights, enhancing prediction accuracy while upholding model interpretability.

2.5.3 Feature Engineering for Spatial-Temporal Data

The analysis of spatial-temporal data presents unique challenges and opportunities within machine learning, particularly for wellbeing monitoring [40–44]. Extracting meaningful features from such dynamic datasets requires specialised techniques capable of capturing intricate spatial and temporal relationships. Recent advancements in feature engineering for spatial-temporal data show promising improvements in predictive accuracy and interpretability across various applications. These advancements offer potential methods applicable to the environmental and physiological data explored in this PhD research programme.

For instance, in predictive maintenance, temporal convolution networks have demonstrated enhanced estimations of remaining useful life [40]. By combining CNN-based spatial feature extraction with LSTM-based temporal dynamics modelling through causal filters and dilated convolutions, this approach efficiently processes degradation signals over long periods. The success highlights the potential for tailored neural networks to extract value from temporal dynamics, which could be adapted for monitoring long-term physiological responses to environmental exposures.

Similarly, graph-based feature extraction techniques, such as the SuperGraph method, provide opportunities to uncover hidden topological insights within spatial-temporal data [41]. Applied to rotational machinery fault diagnosis, SuperGraph transforms raw vibration signals into informative graphs using graph theory principles. Subsequent analysis via Laplacian matrices enables the extraction of discriminative spatial-temporal features for precise classification. While focused on machinery, graph-based methods could potentially model relationships between spatially distributed environmental sensors or complex physiological interactions over time.

On a broader scale, techniques jointly modelling spatial and temporal dependencies improve urban system analysis. Luo et al. [42] developed ESTNet, using multi-scale 3D convolutions

to capture intricate correlations across road networks and real-time traffic data. The comprehensive feature extraction led to enhanced traffic flow predictions, relevant to urban planning aspects of public wellbeing. These approaches for integrating spatial and temporal information are conceptually relevant to analysing how environmental conditions change across space and time and impact individuals.

While advanced feature extraction provides predictive power, dimensionality reduction remains crucial for efficiency and interpretability. Song and Zhao's [43] review highlights the versatility of Slow Feature Analysis (SFA) in emphasising meaningful slowly varying temporal features, applicable in diverse monitoring tasks. Complementary spatial-temporal dimensionality reduction methods also exist; Zhang et al. [44] introduced double-window PCA (DWPCA) for structural health monitoring. Such techniques could be valuable for simplifying the complex environmental and physiological datasets used in wellbeing research.

As the landscape of spatial-temporal feature engineering advances, robust methodologies balancing predictive accuracy and interpretability become increasingly valuable. The innovations explored here not only enhance AI applicability to spatial-temporal tasks but also open possibilities for improving wellbeing analysis through established and emerging domains.

2.5.4 Relevant Machine Learning Algorithms

Machine learning algorithms provide powerful tools for analysing the complex relationships within environmental and physiological data. Based on the nature of the task, different types of algorithms are employed. Supervised learning, using labelled data, encompasses both regression (predicting continuous values) and classification (predicting discrete categories) [29]. Unsupervised learning, dealing with unlabelled data, includes clustering for discovering inherent groupings. This subsection reviews supervised and unsupervised algorithms relevant to environmental wellbeing analysis.

2.5.4.1 Regression Algorithms

Regression algorithms are primarily used for prediction tasks, such as forecasting environmental variables or physiological states based on historical sensor data. Their accuracy is often evaluated by comparing predicted values against actual measurements. Regression models have seen widespread use in environmental modelling due to their interpretability.

Linear Regression models estimate a dependent variable (y) based on a set of independent variables (x), assuming a straight-line relationship [19]. The dependency function is given by:

$$y = Xw + \varepsilon \tag{2.4}$$

Here, $w \in \mathbb{R}^{m+1}$ represents the model parameters (weights), $X \in \mathbb{R}^{n\times(m+1)}$ is the matrix of observations, $\varepsilon \in \mathbb{R}^n$ denotes the error or noise term, and $y \in \mathbb{R}^n$ is the column-vector representing the target variable. Parameters (w) are often determined using methods such as Ordinary Least Squares (OLS) or Gradient Descent (GD). OLS identifies weights that minimise the mean squared error, defined as $L(X, y, w) = \frac{1}{2n}\|y - Xw\|^2 = \frac{1}{2n}(y - Xw)^T(y - Xw)$ [19]. The analytical solution for OLS is:

$$w = (X^T X)^{-1} X^T y \tag{2.5}$$

Alternatively, Gradient Descent iteratively adjusts weights to minimise the loss function L. The update rule for each weight is $w = w - \alpha\frac{\partial}{\partial w}L(X, y, w)$, where α is the learning rate controlling the step size. Despite its simplicity, linear regression remains effective for modeling basic relationships, such as the impact of an environmental pollutant on a health indicator, when such relationships are approximately linear. (see Figure 2.2)

Support Vector Regression (SVR) is a supervised learning technique particularly noted for its use of kernels, sparse solutions, and control over the margin of error [45]. It aims to find a function $f(x, w) = w^T x$ that has at most ε deviation from the actually obtained targets y_i for all the training data, and at the same time, is as flat as possible (minimising $\frac{1}{2}\|w\|^2$). This concept is illustrated in Figure 2.3. Errors are only considered if they lie outside this ε-insensitive tube,

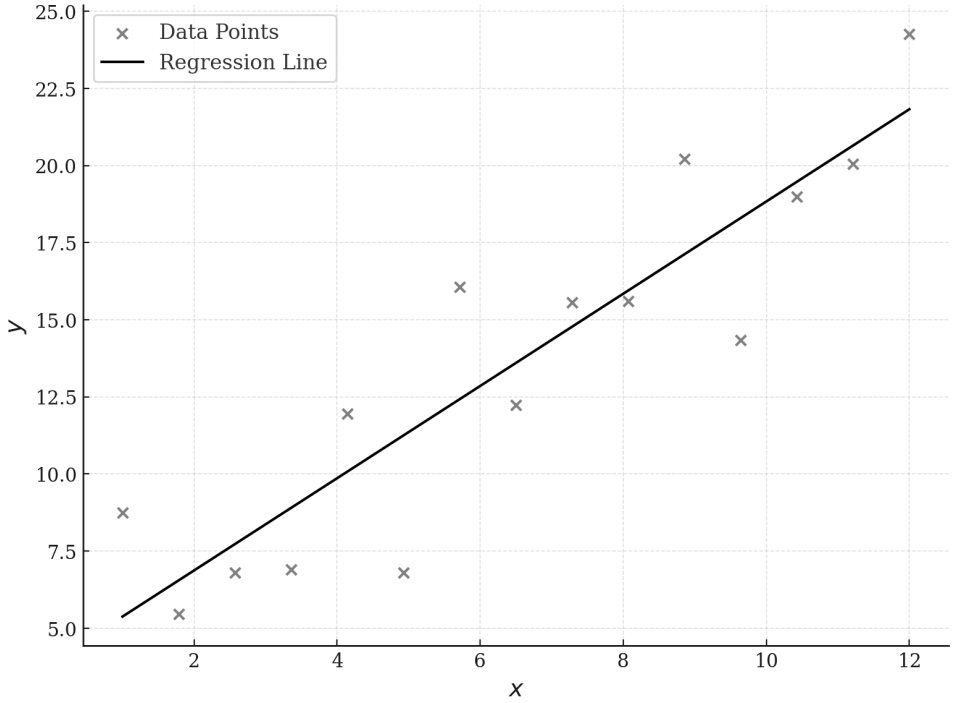


Figure 2.2: Visual Representation of Univariate Linear Regression

defined by the constraint:

$$|y_i - w^T x_i| \leq \varepsilon \tag{2.6}$$

This tolerance makes SVR robust to outliers and useful in high-dimensional spaces, potentially beneficial for modelling complex physiological responses to diverse environmental inputs. Various loss functions can manage errors exceeding the ε margin, adapting the penalty for deviations. Examples include the linear ε-insensitive loss:

$$L_E(y, f(x, w)) = \begin{cases} 0 & \text{if } |y - f(x, w)| \leq \epsilon \\ |y - f(x, w)| - \epsilon & \text{otherwise} \end{cases} \tag{2.7}$$

the quadratic loss:

$$L_E(y, f(x, w)) = \begin{cases} 0 & \text{if } |y - f(x, w)| \leq \epsilon \\ (|y - f(x, w)| - \epsilon)^2 & \text{otherwise} \end{cases} \tag{2.8}$$

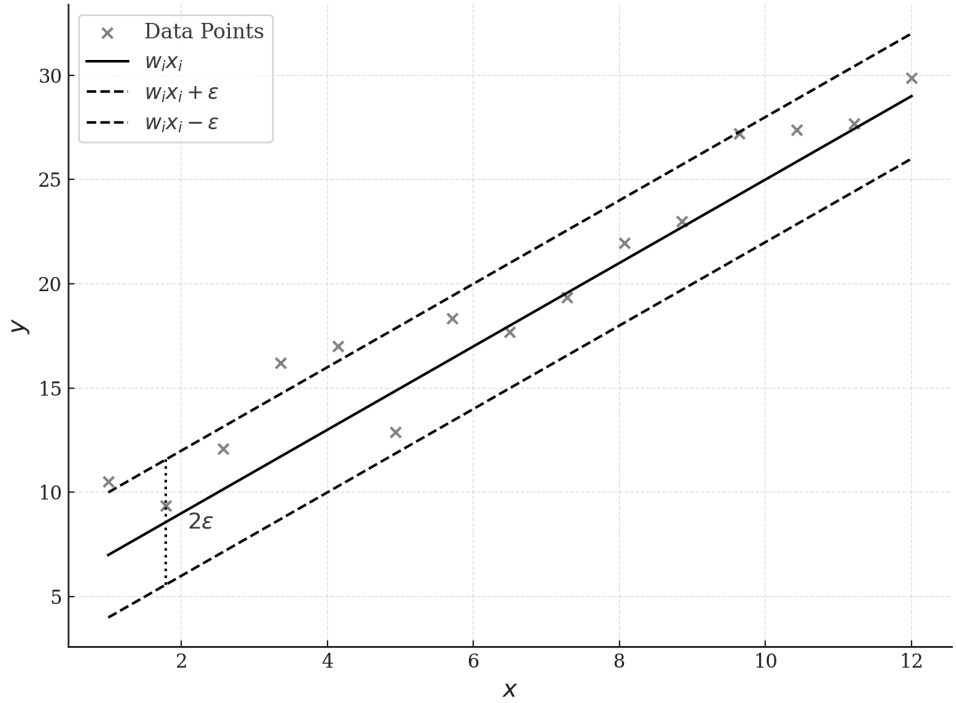


Figure 2.3: Visual Representation of Support Vector Regression

and the Huber loss:

$$L(y, f(x,w)) = \begin{cases} c|y - f(x,w)| - \frac{c^2}{2} & \text{if } |y - f(x,w)| > c\frac{1}{2} \\ \frac{1}{2}|y - f(x,w)|^2 & \text{otherwise} \end{cases} \tag{2.9}$$

These loss functions are visualised in Figure 2.4.

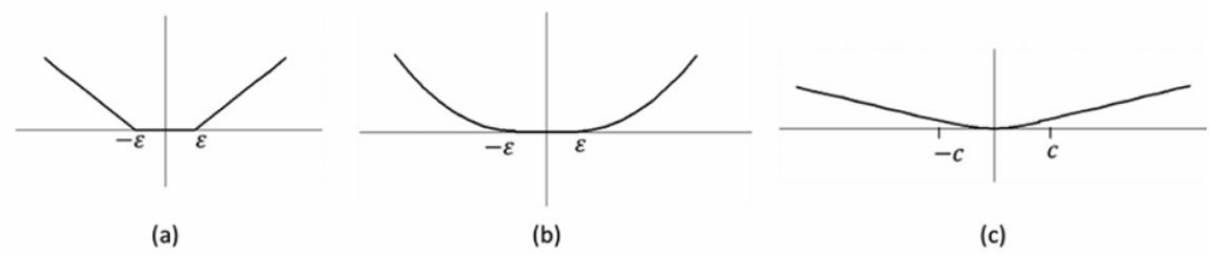


Figure 2.4: SVR Loss function types: (a) linear, (b) quadratic, (c) Huber

k-Nearest Neighbours (k-NN) Regression is a non-parametric method where the prediction for a new input is based on the average of the target values of its k nearest neighbours in the training data [46]. Closeness is typically determined using a distance metric, such as Euclidean

distance:

$$d(x^{[a]}, x^{[b]}) = \sqrt{\sum_{j=1}^{m} (x_j^{[a]} - x_j^{[b]})^2} \tag{2.10}$$

The predicted value $h(x^{(t)})$ for a test point $x^{(t)}$ is then the mean of the target values (y_i) of the k identified neighbours:

$$h(x^{(t)}) = \frac{1}{k} \sum_{i=1}^{k} y_i \tag{2.11}$$

Because k-NN makes predictions based on local proximity rather than a global model, it can effectively capture complex, non-linear relationships in heterogeneous data distributions.

Gradient Boosting (GB) Regression constructs a strong predictive model by sequentially adding weak learners, typically decision trees, where each new learner focuses on correcting the errors made by the previous ones [47]. This ensemble method optimises a differentiable loss function $L(y, F(x))$ via a gradient descent procedure in function space. Starting with an initial model $F_0(x)$, the algorithm iterates from $m = 1$ to M. In each iteration, it computes pseudo-residuals r_{im} (the negative gradient of the loss function with respect to the model output $F(x)$ evaluated at $F_{m-1}(x)$). A base learner $h_m(x)$ is then fitted to these pseudo-residuals. An optimal multiplier γ_m is computed, often through a line search, to determine the best step size for the new learner. Finally, the model is updated: $F_m(x) = F_{m-1}(x) + \gamma_m h_m(x)$ [47]. GB methods are known for high accuracy and are well-suited for complex prediction tasks involving numerous environmental and physiological features.

Gaussian Process Regression (GPR) employs a non-parametric, Bayesian approach [48]. Instead of learning specific parameters of a function, GPR defines a prior distribution over possible functions. When provided with training data $(X, \mathbf{y})$, this prior is updated using Bayes' theorem to obtain a posterior distribution over functions that are consistent with the observed data [48]. Predictions for new inputs X_* are derived from this posterior distribution, yielding both a mean prediction $\mathbf{f}_*$ and a measure of uncertainty (variance) around that prediction [48]. The Gaussian process prior is specified by a mean function $m(x)$ and a covariance function (kernel) $k(x, x')$, such that $f(x) \sim \text{GP}(m(x), k(x, x'))$. Assuming independent Gaussian noise $\varepsilon \sim N(0, \sigma_n^2)$ on the observations ($y = f(x) + \varepsilon$), the joint distribution of the observed targets $\mathbf{y}$ and the function

values $\mathbf{f}_*$ at the test points X_* is multivariate Gaussian:

$$\begin{pmatrix} \mathbf{y} \\ \mathbf{f}_* \end{pmatrix} \sim \mathcal{N}\left(\begin{pmatrix} \mu \\ \mu_* \end{pmatrix}, \begin{bmatrix} K(X,X) + \sigma_n^2 I & K(X,X_*) \\ K(X_*,X) & K(X_*,X_*) \end{bmatrix} \right) \tag{2.12}$$

Here, K denotes the covariance matrix obtained by evaluating the kernel function k at the respective pairs of data points [48]. GPR is particularly useful for smaller datasets or when quantifying prediction uncertainty is crucial, such as assessing risks associated with environmental exposures. Figure 2.5 illustrates the concept of updating the prior to a posterior distribution.

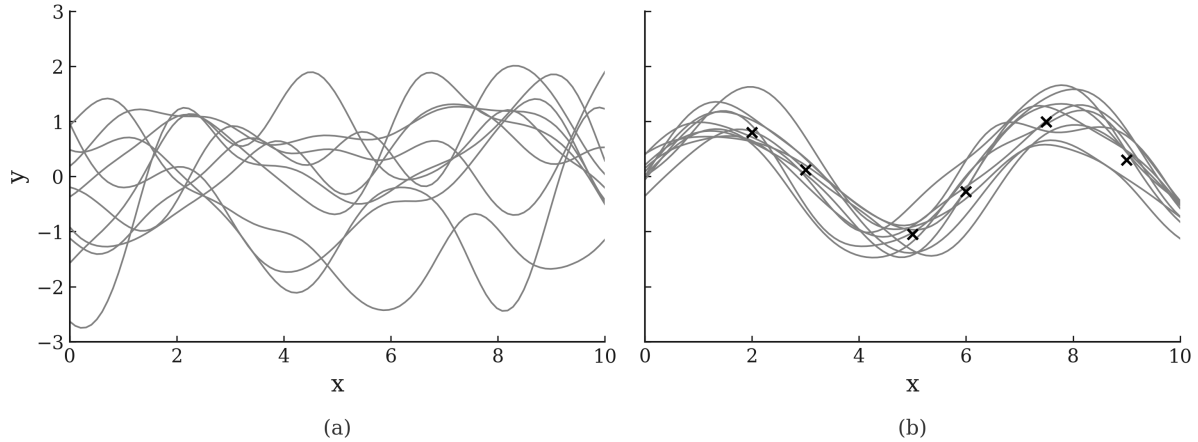


Figure 2.5: Gaussian Process Regression: (a) Prior and (b) Posterior

2.5.4.2 Classification Algorithms

Classification algorithms, another category of supervised learning, are used to assign data points to predefined categories or classes. These are essential in environmental wellbeing research for tasks such as identifying environmental states (e.g., polluted vs. unpolluted) or classifying physiological states (e.g., stressed vs. relaxed) based on sensor data.

Random Forest (RF) and Decision Trees (DT). A Random Forest operates as an ensemble of multiple Decision Trees [49]. Each individual tree in the forest provides a class prediction, and the final output is determined by a majority vote among all trees (Figure 2.6). This ensemble approach generally improves accuracy and robustness compared to a single tree. A Decision Tree itself models decisions in a tree-like structure (Figure 2.7). Internal nodes represent tests on attributes (features), branches represent the outcome of the test, and leaf nodes represent

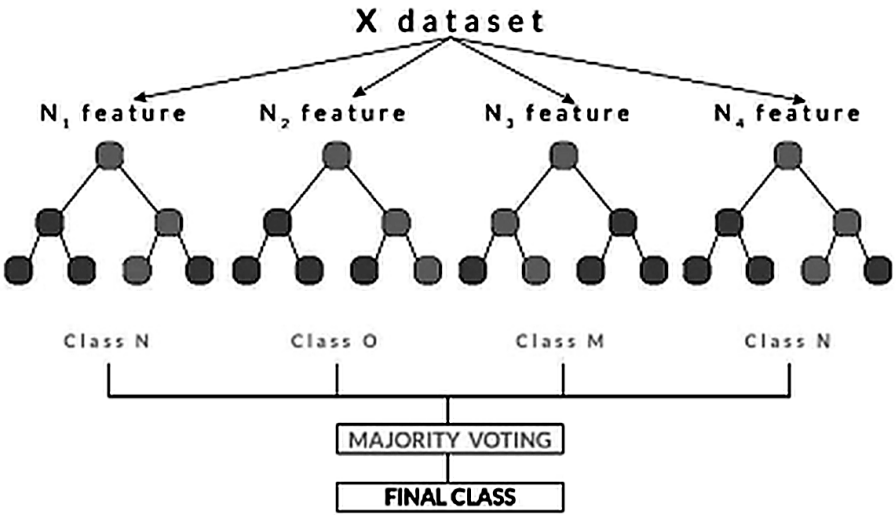


Figure 2.6: Visualization of a Random Forest Model Prediction

the final class label [50]. The tree construction process typically involves recursively selecting the best attribute to split the data at each node, aiming to purify the resulting subsets. This selection is often guided by Attribute Selection Measures (ASM) like Information Gain, which uses entropy to quantify the reduction in uncertainty achieved by a split. Shannon's entropy for a system with N states is defined as $S = -\sum_{i=1}^{N} p_i \log_2 p_i$, where p_i is the probability of state i. Information Gain $IG(Q)$ for a split on attribute Q is calculated as the entropy before the split (S_0) minus the weighted average entropy of the subsets after the split: $IG(Q) = S_0 - \sum_{i=1}^{q} \frac{N_i}{N} S_i$, where q is the number of groups after the split, N_i is the number of instances in group i, and S_i is the entropy of group i. The splitting process continues until a stopping criterion is met, such as all instances belonging to one class or no further attributes remaining. Decision trees offer interpretability, mapping clear decision pathways, which can be valuable for understanding environmental factors influencing wellbeing classifications [51].

Bagging (Bootstrap Aggregating). This is the ensemble technique underlying Random Forests [50]. It involves creating multiple bootstrap samples (random samples with replacement) from the original training set (D). For each bootstrap sample (D_i'), a base model (e.g., a decision tree) is trained. The final prediction is obtained by aggregating the predictions of all base models, typically through majority voting for classification. Bagging helps reduce variance and improve the stability of the models.

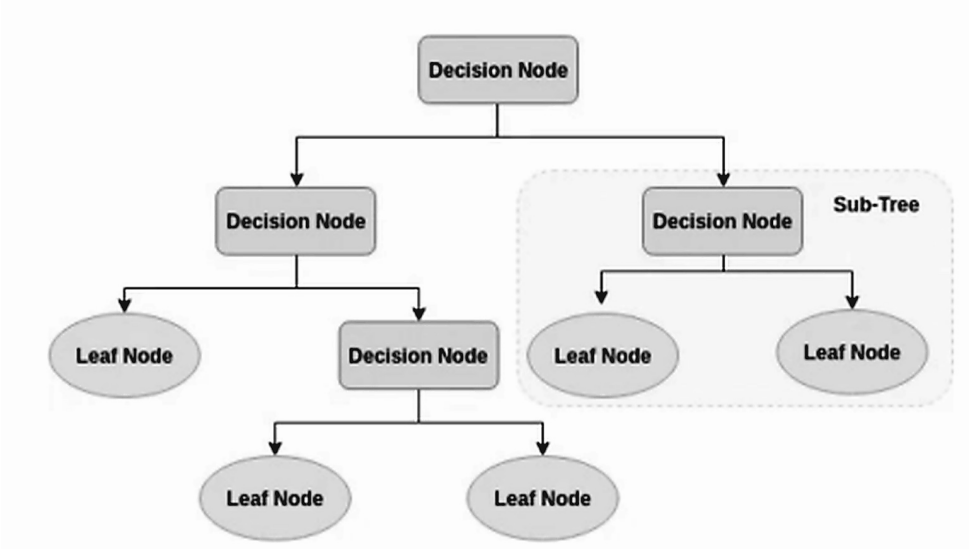


Figure 2.7: Visualization of a Decision Tree Model

Artificial Neural Networks (ANN). ANNs are supervised learning algorithms inspired by biological neural networks [48]. They consist of interconnected nodes or neurons organised in layers (input, hidden, output). Each neuron performs a computation, typically a weighted sum of its inputs followed by a non-linear activation function (see Figure 2.9 for common activation functions), and passes the result to neurons in the next layer. A network with only one hidden layer is considered "shallow," while a Deep Neural Network (DNN) has multiple hidden layers (Figure 2.8), allowing it to model more complex, non-linear relationships often found in environmental and physiological data. Key concepts in ANNs include inputs (data fed to the network), outputs (predictions), neurons (basic processing units), activation functions (introducing non-linearity), an error or loss function (measuring prediction discrepancy), backpropagation (algorithm for adjusting weights to minimize error), and hyperparameters (tunable settings like layer count, learning rate, epochs).

Convolutional Neural Networks (CNN). CNNs are a type of DNN particularly effective for processing grid-like data, such as images or time series [52]. A typical CNN architecture includes convolutional layers, pooling layers, and fully connected layers (Figure 2.10). Convolutional layers apply filters (kernels) across the input data to detect spatial or temporal patterns, creating feature maps. Pooling layers then reduce the dimensionality of these feature maps (e.g., max pooling), retaining essential information while reducing computational load. Finally, fully connected layers integrate the learned features to perform classification. CNNs are adept

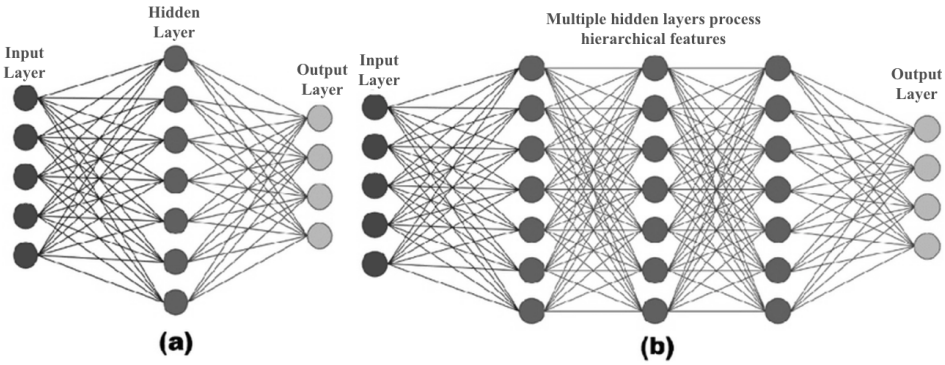


Figure 2.8: Feed Forward Neural Network Architectures: (a) Shallow, (b) Deep)

at capturing hierarchical patterns automatically.

While often used for 2D data (images), 1D-CNNs are adapted for sequential data like time series from sensors (Figure 2.11) [53–55]. They apply 1D filters along the time axis, making them suitable for identifying temporal patterns in physiological or environmental sensor readings relevant to wellbeing analysis.

Polysomnography (PSG) involves recording multiple physiological signals during sleep. Machine learning models leverage this rich data to uncover patterns related to sleep quality and disorders, which significantly impact wellbeing. By analysing PSG data, ML algorithms such as Logistic Regression (LR), k-NN, and SVM can help detect conditions like sleep apnea or insomnia [56,57] or classify states related to mental health [58,59]. Table 2.1 summarises several studies applying these machine learning algorithms to PSG data analysis.

Table 2.1: A summary of ML-based studies for classification and prediction of PSG data. Dataset sources are cited within the table.

Model	Application	Data Used	Accuracy	Year Ref.
LR	EEG abnormalities in temporal lobe epilepsy (TLE)	Privately sourced dataset from tertiary institute	66.70%	2018 [58]
LR	Mental depression detection	*emotions.csv*, Kaggle	96.60%	2022 [59]

Model	Application	Data Used	Accuracy	Year Ref.
LR	Emotion Recognition	DREAMER Dataset [60]	94.20%	2021 [61]
k-NN	ECG-based stress/emotion classification	Not specified	97.00%	2022 [62]
k-NN	OSA using ECG and SpO2	PhysioNet Sleep Apnea Database [63]	95.08%	2017 [56]
SVM	EEG emotion classification	SEED Dataset [64]	56.00%	2022 [65]
SVM	OSA using ECG and SpO2	PhysioNet Sleep Apnea Database [63]	96.64%	2017 [56]
SVM	Sleep disorder classification	Sleep-EDF Database [66]	91.40%	2019 [57]
SVM	ADHD detection using imaging + EEG	ADHD-200 Dataset	97.60%	2022 [67]
SVM	Human identification using EEG	EMOTIV INSIGHT Dataset [68]	94.44%	2016 [68]
SVM	Mental stress detection	EEG Mental Arithmetic Dataset [69]	97.26%	2022 [69]
SVM	EEG dimensionality reduction	BCI Competition II - Dataset III [70]	81.40%	2017 [71]
SVM	Motor imagery classification	BCI Competition II - Dataset III [70]	78.57%	2019 [72]
SVM	Alcoholism detection via ECG	NIMHANS ECG Dataset	87.50%	2017 [73]

Model	Application	Data Used	Accuracy	Year Ref.
SVM	Sleep quality measurement	Sleep-EDF Database [66]	93.50%	2019 [57]
SVM	Mental depression via EEG	*emotions.csv*, Kaggle	95.89%	2022 [59]
SVM	Schizophrenia detection	EEG dataset from M.V. Lomonosov Moscow State University [74]	53.50%	2022 [74]

Application to Environmental Sensor Data. While the classification algorithms discussed above are broadly applicable, their utility becomes particularly evident in environmental monitoring scenarios. For instance, Random Forest classifiers have been widely adopted for classifying air quality levels based on multivariate sensor inputs, such as particulate matter (PM2.5), NO_2, and CO concentrations [75]. Similarly, Support Vector Machines have been successfully deployed in water quality assessment systems to distinguish between contaminated and safe samples using features like pH, turbidity, and electrical conductivity [46]. In real-time anomaly detection tasks such as identifying abnormal patterns in temperature or humidity readings from wireless sensor networks k-NN and Decision Trees are frequently used due to their interpretability and ease of implementation. Moreover, Convolutional Neural Networks (CNNs), although traditionally used for image data, have recently been repurposed to classify time-series data from wearable or fixed environmental sensors by transforming sequential inputs into 2D feature maps [76]. These tailored implementations demonstrate that machine learning and deep learning techniques offer not just theoretical flexibility, but concrete, scalable solutions for classifying environmental conditions with high precision.

2.5.4.3 Clustering and Dimensionality Reduction Algorithms

Unsupervised learning methods, such as clustering and dimensionality reduction, are valuable for exploring environmental and physiological data to uncover hidden structures or simplify

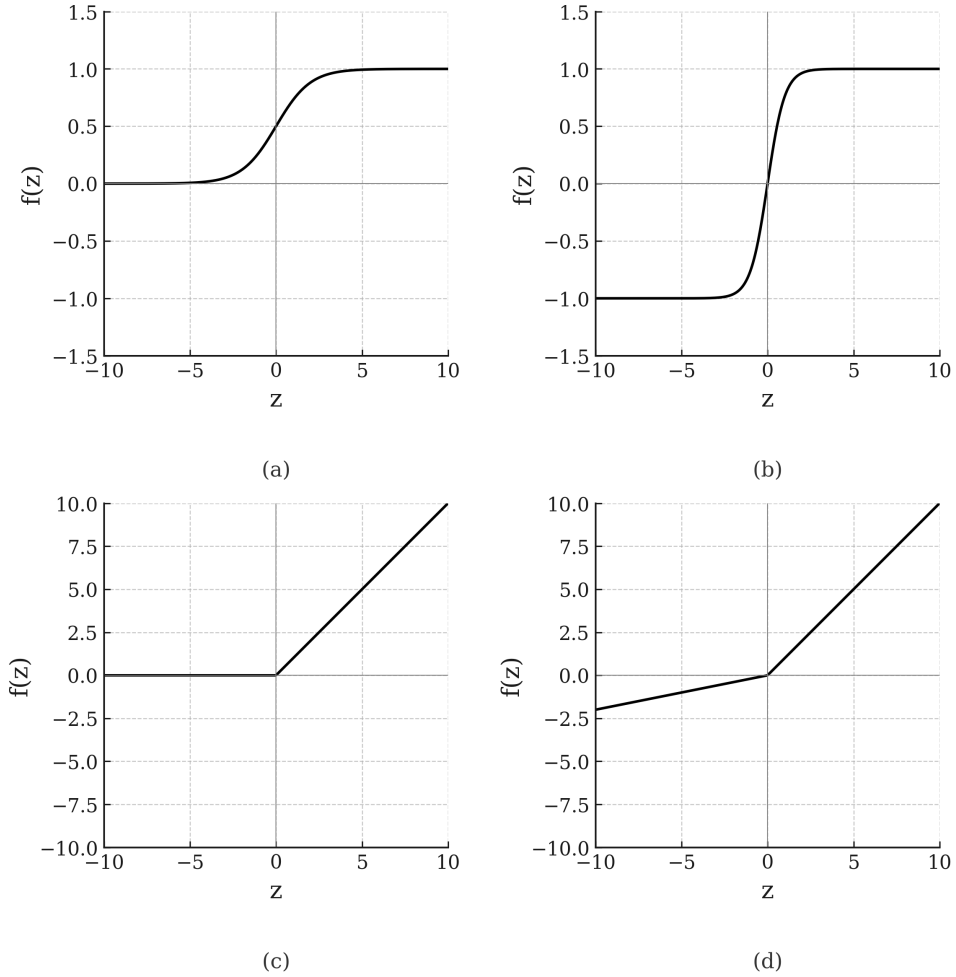


Figure 2.9: Common Activations Functions: (a) Sigmoid, (b) Tanh, (c) ReLU, (d) LeakyReLU

complexity without predefined labels.

Principal Component Analysis (PCA). PCA is a widely used technique for dimensionality reduction. Its goal is to transform a dataset containing many potentially correlated variables into a smaller set of uncorrelated variables, known as principal components, while preserving as much of the original dataset's variance as possible [77]. This technique is particularly useful when dealing with high-dimensional sensor data, common in environmental wellbeing studies, as it helps reduce computational complexity and potentially filter noise. The process generally involves standardising the data, computing the covariance matrix, calculating its eigenvectors and corresponding eigenvalues, sorting the eigenvectors based on the magnitude of the eigenvalues (which represent the variance captured by each component), and finally projecting the original data onto the subspace spanned by the selected principal components (typically those capturing the most variance).

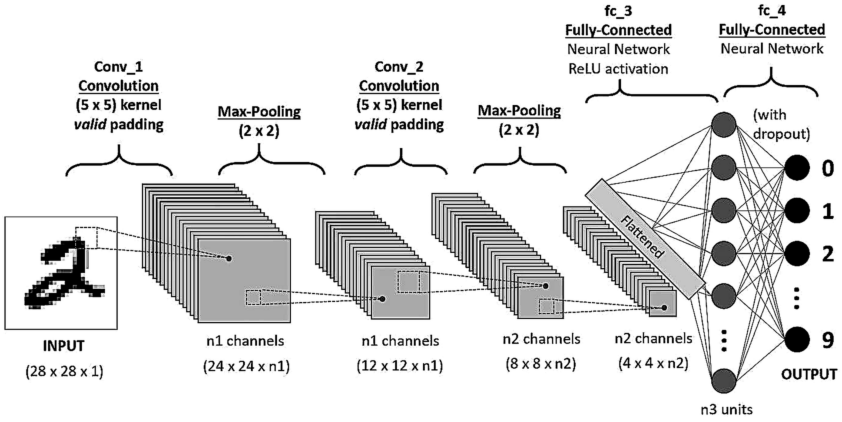


Figure 2.10: Convolutional Neural Network Architecture for Handwritten Digit Recognition

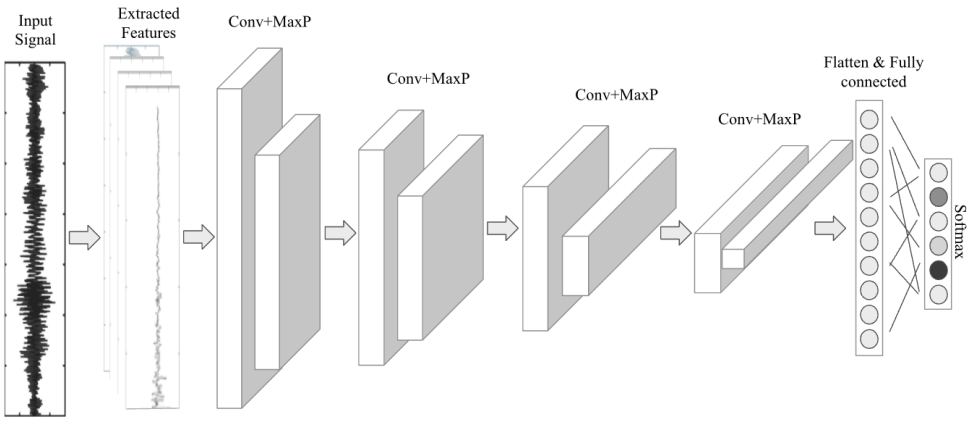


Figure 2.11: Illustration of 1D-CNN architecture

K-means Clustering. K-means is a widely used unsupervised algorithm that partitions a dataset into a specified number (k) of distinct, non-overlapping clusters [78]. The algorithm works iteratively: it first randomly initialises k centroids, then assigns each data point to the nearest centroid (commonly based on Euclidean distance), and subsequently recalculates each centroid as the mean of all points assigned to it. These assignment and update steps are repeated until the centroids no longer move significantly or a maximum number of iterations is reached. In the context of environmental wellbeing research, K-means could potentially be applied to identify distinct environmental condition profiles based on multivariate sensor readings or to group study participants based on similarities in their physiological response patterns under different conditions.

Autoencoder. An autoencoder is a type of ANN utilised for unsupervised learning, often employed for dimensionality reduction or learning efficient data codings [48]. Structurally, as

depicted in Figure 2.12, it comprises two main components: an encoder and a decoder. The encoder network maps the input data into a lower-dimensional latent space, creating a compressed representation or encoding that ideally captures the essential features of the input. The decoder network then attempts to reconstruct the original input data accurately from this latent representation. The autoencoder is trained by minimising the reconstruction error (the difference between the original input and the reconstructed output). Unlike PCA which performs a linear transformation, autoencoders can learn complex, non-linear mappings, making them potentially suitable for compressing non-linear physiological or environmental time-series data while retaining important underlying patterns.

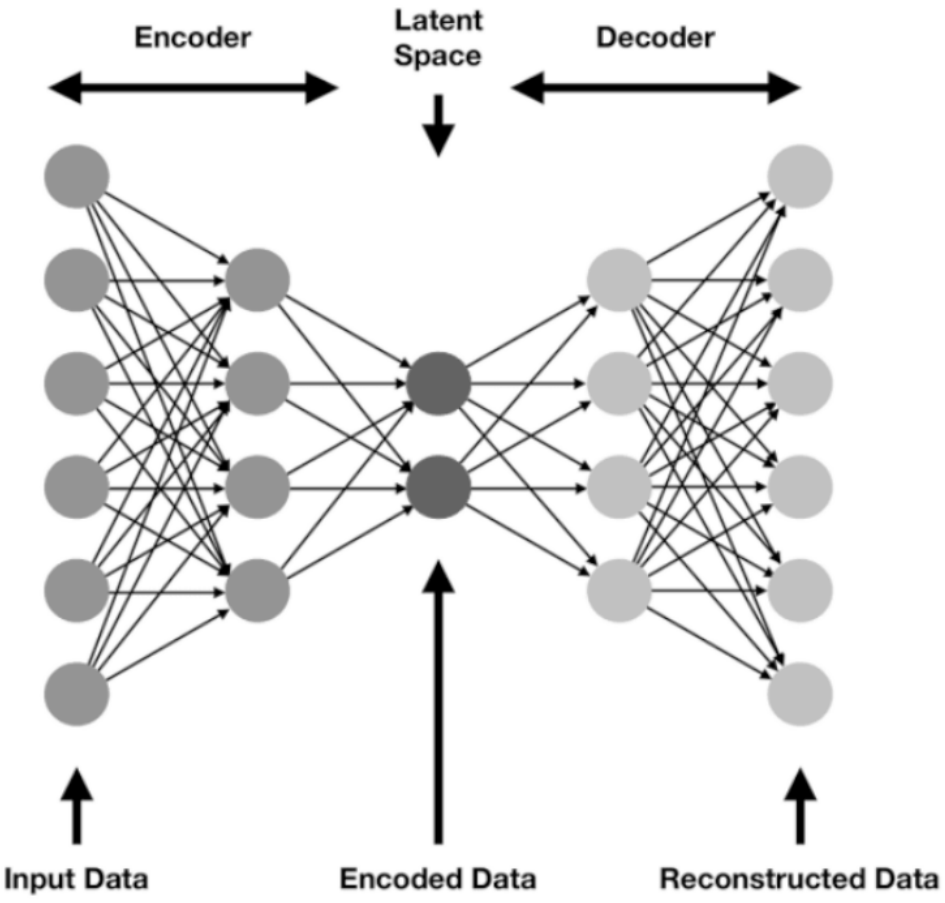


Figure 2.12: Illustration of simple Autoencoder Architecture

2.5.4.4 Deep Learning Algorithms

Deep learning offers powerful tools for modelling complex patterns, especially in sequential and spatial-temporal data common in environmental and physiological monitoring. This subsection

reviews key deep learning architectures relevant to this research programme.

Recurrent Neural Networks (RNN). RNNs are designed to process sequential data by maintaining an internal memory or hidden state (H) that captures information from previous time steps [19]. At each time step t, the RNN takes the previous hidden state $H^{<t-1>}$ and the current input $X^{<t>}$ to compute the new hidden state $H^{<t>}$ and an output $\hat{Y}^{<t>}$ (Figure 2.13). The core equations are:

$$H^{<t>} = g(H^{<t-1>}W_{hh} + X^{<t>}W_{hx} + \mathbf{1}\mathbf{b}_{\mathbf{h}}^{\top}) \tag{2.13}$$

$$\hat{Y}^{<t>} = h(H^{<t>}W_{yh} + \mathbf{1}\mathbf{b}_{\mathbf{y}}^{\top}) \tag{2.14}$$

where $W_{hh}, W_{hx}, W_{yh}, \mathbf{b_h}, \mathbf{b_y}$ are learnable weights and biases, and g, h are activation functions [19]. This architecture allows RNNs to handle variable-length sequences and model temporal dependencies. However, simple RNNs often struggle with learning long-range dependencies due to the vanishing gradient problem [79].

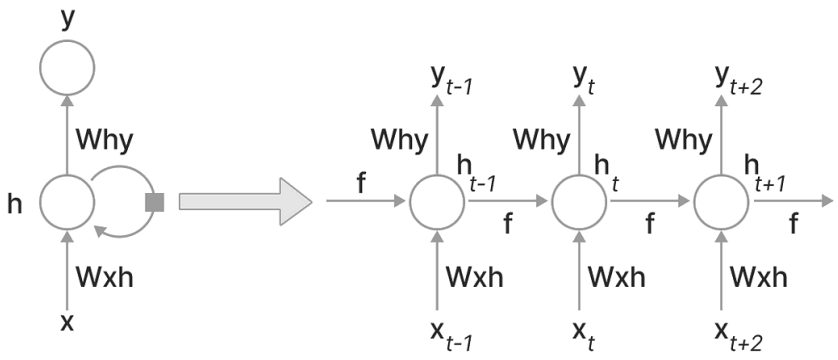


Figure 2.13: The architecture of an RNN.

Long Short-Term Memory (LSTM) and Gated Recurrent Units (GRU). To address the limitations of simple RNNs, more complex recurrent units like LSTM [80] and GRU [81] were developed. LSTMs introduce a memory cell (C) and gating mechanisms (input, forget, output gates $\Gamma_u, \Gamma_f, \Gamma_o$) that control the flow of information, allowing the network to selectively remember or forget information over long sequences. The key LSTM equations are:

$$\Gamma_u = \sigma(W_u[H^{<t-1>}, X^{<t>}] + \mathbf{b_u}\mathbf{1}^{\top}) \tag{2.15}$$

$$\Gamma_f = \sigma(W_f[H^{<t-1>}, X^{<t>}] + \mathbf{b_f}\mathbf{1}^\top) \tag{2.16}$$

$$\Gamma_o = \sigma(W_o[H^{<t-1>}, X^{<t>}] + \mathbf{b_o}\mathbf{1}^\top) \tag{2.17}$$

$$\tilde{C}^{<t>} = \tanh\left(W_c[H^{<t-1>}, X^{<t>}] + \mathbf{b_c}\mathbf{1}^\top\right) \tag{2.18}$$

$$C^{<t>} = \Gamma_u \odot \tilde{C}^{<t>} + \Gamma_f \odot C^{<t-1>} \tag{2.19}$$

$$H^{<t>} = \Gamma_o \odot \tanh(C^{<t>}) \tag{2.20}$$

where σ is the sigmoid function, $\tanh$ is the hyperbolic tangent, $\odot$ denotes element-wise multiplication, and $[H^{<t-1>}, X^{<t>}]$ represents concatenation. GRUs offer a simpler alternative with fewer parameters, merging the input and forget gates into a single update gate, often achieving comparable performance. Both LSTMs and GRUs have proven effective in analysing environmental time series, such as forecasting air quality or temperature, by capturing complex temporal patterns relevant to public health [82].

Specialized Time-Series Transformers for Environmental and Physiological Applications. Transformer architectures, originally introduced for natural language processing [83], have achieved significant success in time-series modeling due to their ability to efficiently capture long-range temporal dependencies through self-attention mechanisms [84, 85]. This modeling capability has led to widespread adoption in domains such as environmental forecasting and physiological signal processing.

Recent developments have produced architectures tailored to the unique challenges of time-series forecasting. The Temporal Fusion Transformer (TFT) [86] combines recurrent layers, attention mechanisms, and gating components to handle both static and time-varying inputs, and has demonstrated strong performance in applications like air quality forecasting, energy consumption modeling, and climate prediction. The Informer [84] introduces a ProbSparse self-attention mechanism to improve computational efficiency, enabling effective handling of high-resolution, long-sequence data in meteorological forecasting and pollutant dispersion simulations.

In physiological signal analysis, Transformer-based models continue to set new benchmarks. Lu

et al. proposed PainAttnNet [87], a hybrid model that combines multiscale convolutional layers with a Transformer encoder for classifying pain intensity from biosignals, achieving higher accuracy than conventional models. Similarly, Wang et al. introduced PhysioFormer [88], which integrates Transformer learning with symbolic regression to predict affective states from multimodal physiological signals. Tested on the WESAD dataset [89], PhysioFormer achieved over 99% accuracy, highlighting both the robustness and explainability of Transformer-based approaches for health monitoring.

Integrated Spatial-Temporal Deep Learning Models. For data with both spatial and temporal dimensions, specialised architectures like 3D CNNs and Convolutional LSTMs (ConvLSTMs) are used [90–93]. 3D CNNs extend 2D CNNs by applying convolutions across the time dimension as well, useful for capturing short-term spatio-temporal features, for instance in video analysis or near-term traffic prediction [90]. ConvLSTMs combine convolutional operations within the LSTM cell structure, allowing them to learn spatial patterns and temporal dynamics simultaneously. They are particularly suited for tasks like precipitation nowcasting or predicting air quality distribution over time based on sensor networks [92]. Blended architectures, like combining CNN feature extraction with Bi-directional LSTMs (Conv-BiLSTM), have also shown high accuracy in spatio-temporal forecasting tasks such as traffic congestion prediction [91]. Table 2.2 provides a summary of deep learning studies applied to Polysomnography (PSG) data, illustrating applications in health-related time series analysis.

These integrated models enable robust spatial-temporal analysis, with ConvLSTMs particularly strong for long sequences, driving innovation in real-time monitoring and predictive analytics.

2.5.4.5 Time Series Specific Techniques

Classifying time series data, such as environmental sensor readings or physiological signals over time, often requires specialised techniques beyond standard classification algorithms applied to static feature vectors. Key considerations include handling variable lengths and capturing temporal dependencies.

Dynamic Time Warping (DTW). When comparing time series, especially those that may be

Table 2.2: A summary of deep learning-based studies for classification and prediction of PSG data.

Model	Application	Data Used	Accuracy	Year Ref.
ResNet-50	EEG image data and emotion classification	SEED Dataset	85.11%	2022 [65]
CNN	EEG-sleep stage using multi-scale dual-attention	Sleep-EDF Database	96.70%	2022 [94]
	Mental depression from EEG dataset	emotions.csv available on the Kaggle website	49.82%	2022 [59]
	Automatic sleep scoring	Multiple EGG dataset was used for this work	74.17%	2021 [95]
	Emotion recognition	DREAMER (discrete emotion recognition)	99.90%	2021 [61]
ELM	Identification of chronic alcohol users from ECG signals	NIMHANS-ECG Dataset	94.64%	2017 [73]
MLP	Mental depression from EEG dataset	emotions.csv available on the Kaggle website	76.43%	2022 [59]
RNN	Mental depression from EEG dataset	emotions.csv available on the Kaggle website	93.90%	2022 [59]
RNN with LSTM	Mental depression from EEG dataset	emotions.csv available on the Kaggle website	97.65%	2022 [59]
	Detection of schizophrenia from EEG Data	EEG dataset from NNCI M. V. Lomonosov Moscow State University	98.00%	2022 [74]
	Insomnia detection	MASS Dataset-EEG, EOG, EMG, ECG, and respiratory signals	79.60%	2021 [96]
	Depression using EEG	BCI project for EEG signal and frontal facial data	99.66%	2021 [97]
CNN-LSTM	Automatic sleep scoring	Multiple EGG dataset was used for this work	80.17%	2021 [95]
	Sleep apnea	Apnea-ECG dataset	97.21%	2022 [98]

out of phase or of different lengths, standard distance metrics like Euclidean distance can be misleading. DTW is a widely used algorithm that finds the optimal non-linear alignment between two time series [99]. It calculates a distance measure that allows similar patterns to be matched even if they occur at slightly different times or speeds [99]. This makes DTW particu-

larly suitable for comparing physiological responses or environmental patterns that may exhibit temporal variability.

DTW with k-Nearest Neighbours (DTW-KNN). The k-Nearest Neighbours (KNN) algorithm can be adapted for time series classification by using DTW as the distance metric [46]. Instead of finding the k neighbours based on Euclidean distance in a feature space, DTW-KNN finds the k training time series that are most similar to a new, unlabelled time series according to the DTW distance. The new series is then classified based on a majority vote among these k nearest neighbours (Figure 2.14). This approach leverages the pattern-matching strength of DTW within the simple framework of KNN.

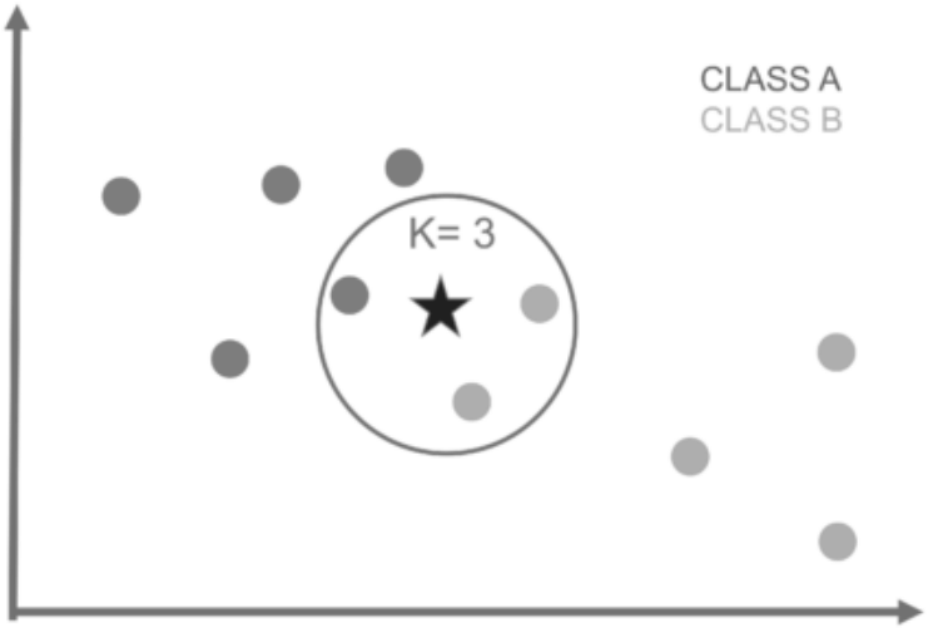


Figure 2.14: KNN algorithm concept (adapted for time series with DTW distance).

Support Vector Machines (SVM) with Time Series Kernels. Support Vector Machines (SVMs) are robust supervised learning models that aim to find an optimal hyperplane for separating data into classes by maximizing the margin between support vectors [45]. When dealing with non-linearly separable data, the SVM leverages the kernel trick to implicitly project the data into a higher-dimensional feature space where a linear separation may become feasible. A commonly used kernel in this context is the Radial Basis Function (RBF) kernel, defined as:

$$k(\vec{x_i}, \vec{x_j}) = \exp\left(-\gamma \|\vec{x_i} - \vec{x_j}\|^2\right),$$

which enables SVMs to learn complex nonlinear decision boundaries [100].

While traditional SVMs operate on fixed-length feature vectors, their application to time series requires either careful feature engineering or the design of sequence-aware kernels. A common strategy involves extracting meaningful statistical features or using dimensionality reduction techniques such as PCA to compress temporal information. Alternatively, Autoencoders can be employed to encode temporal dynamics into compact embeddings that serve as inputs to SVMs. For more direct time series classification, custom kernels based on Dynamic Time Warping (DTW) or other shape-aware similarities can be used to preserve temporal structure [101–104].

The standard soft-margin SVM objective remains:

$$\min_{\vec{w},\xi,b} \frac{1}{2}\|\vec{w}\|^2 + C\sum_{n=1}^{N} \xi_n \quad \text{subject to} \quad y_n(\vec{w}^\intercal \phi(\vec{x_n}) + b) \geq 1 - \xi_n, \quad \xi_n \geq 0,$$

where $\phi(\cdot)$ denotes the implicit feature mapping defined by the kernel function, and $k(\vec{x_i}, \vec{x_j}) = \phi(\vec{x_i})^\intercal \phi(\vec{x_j})$. By selecting or designing kernels that respect the structure of time series, SVMs can effectively classify complex environmental or physiological temporal patterns.

Table 2.3 offers a concise comparison of major machine learning and deep learning algorithms used in environmental and wellbeing analysis. By outlining typical use cases, advantages, limitations, and data suitability, the table assists researchers in selecting appropriate analytical methods for various environmental and physiological datasets.

Table 2.3: Comparative Overview of Machine Learning Algorithms for Environmental and Wellbeing Analysis

Algorithm	Type	Common Use Cases	Advantages	Limitations	Typical Data Types	Suitability
Linear Regression	Regression	Forecasting environmental trends [19, 50, 105]	Interpretable, efficient	Assumes linearity, sensitive to outliers	Numeric, tabular, time-series	Medium
Support Vector Regression	Regression	Pollution forecasting, physiological modeling [45, 100, 106]	Robust, effective in high-dimensional spaces	Computational complexity, kernel selection	Numeric, tabular, time-series	Medium-high
Gaussian Process Regression	Regression	Risk assessment, prediction [48]	Uncertainty quantification, small dataset robustness	High computational cost	Numeric, spatial, time-series	Medium-high
Random Forest	Classification	Air quality, stress detection [49, 107]	Robust, handles complexity, feature importance	Lower interpretability, intensive	Multivariate, spatial, mixed data	High
k-Nearest Neighbours	Classification	Pollution, anomaly detection [46]	Simple, nonlinear data handling	Sensitive to dimensionality, expensive at inference	Numeric, spatial, temporal	Medium
Gradient Boosting	Classification	Air quality, stress responses [47]	High accuracy, robust	Computational complexity, tuning needed	Multivariate tabular data	High
PCA	Dimensionality Reduction	Reducing sensor data complexity [71, 77, 102]	Noise removal, interpretability	Linear assumption, information loss risk	Numeric, high-dimensional data	Medium-high
Autoencoders	Dimensionality Reduction	Sensor data compression [48]	Automatic, non-linear feature learning	Requires large datasets, interpretability issues	Numeric, physiological signals	High
K-means Clustering	Clustering	Environmental zoning, pattern discovery [78, 108, 109]	Simple, exploratory analysis	Requires cluster count, initialization sensitive	Multivariate numeric, spatial	Medium
CNNs	Deep Learning	Remote sensing, physiological imaging [52–55, 65, 110, 111]	Strong pattern recognition	High computational cost, low interpretability	Spatial imagery, sensor data	High
LSTMs/RNNs	Deep Learning	Forecasting, physiological analysis [64, 74, 79, 80, 82, 112]	Captures temporal dependencies	Overfitting risk, computationally demanding	Time-series, physiological data	High
Transformers	Deep Learning	Climate modeling, signal analysis [83–86, 88]	Long-range dependencies, high performance	High computational demands, complex training	Complex time-series, multimodal data	High

2.6 Applications and Case Studies

2.6.1 Urban Wellbeing

Urban forests, encompassing parks, gardens, woodlands, and street trees within urban areas, have become increasingly recognised in recent years for their vital contributions to public mental health and wellbeing [113, 114]. Multiple studies have explored and highlighted the positive impacts of access to biodiverse, well-maintained urban forests on reducing stress, anxiety, and depression, as well as enhancing overall psychological wellbeing, for urban residents.

An analysis by Lafantasie [115] utilising eBird citizen science data suggested that urban green spaces with higher biodiversity and species richness were associated with improved mental health outcomes. Similarly, a longitudinal UK study revealed adolescents with higher daily exposure specifically to urban woodlands had better cognitive development and lower risks of emotional and behavioural disorders compared to those exposed primarily to other urban green space types [116]. These findings emphasise the particularly significant mental health benefits offered by biodiverse, woodland ecosystems within cities.

Beyond observational studies, an urban forest therapy randomised controlled trial (RCT) conducted by Yeon et al. [117] provides further evidence for the positive psychological impacts of urban forests. Participants in forest therapy programmes showed significant reductions in depression, improved sleep quality, and alleviation of stress-linked somatic symptoms. Such empirical evidence substantiates urban forests as an effective, non-pharmacological nature-based solution for depression.

Incorporating psychometric analytics, Liu et al. [118] performed facial expression analysis of visitors to urban forest parks in Northern China. Exposure to the urban forest environments frequently elicited increased positive emotions, indicating enhanced psychological wellbeing.

Collectively, the research underscores urban forests as invaluable city infrastructure for augmenting mental capital and mitigating the psychological toll of urbanisation. Integrating and maintaining urban forests emerges as an imperative public health strategy for city planning au-

thorities. The creation of accessible, biodiverse and well-maintained urban forests can cater to the rising mental health challenges faced by urban populaces. As key hotspots for human-nature interactions, urban forests offer a natural, non-pharmacological and low-cost pathway to managing escalating rates of urban stress, anxiety and depression. Their further incorporation into cities presents immense scope for ecologically conscious urban development that promotes holistic population mental health and wellness.

2.6.2 Green spaces

The relationship between access to urban green spaces and mental health outcomes has garnered increasing research attention in recent years. A growing body of evidence underscores the mental health benefits of nature exposure and interactions for urban populations across diverse geographical contexts. This subsection synthesises key findings from recent studies, situating green spaces as a crucial component of healthy city planning and development.

Multiple large-scale analyses utilising objective green space measures and mental health data provide compelling evidence linking urban nature access with enhanced psychological well-being. Geneshka et al. [119] leveraged European urban green and blue space data integrated with the UK Biobank study, revealing positive associations between the availability of parks, trees, and water bodies with better mental health indicators. At a macro scale, their research highlights the manifold health-promoting capacities of these green and blue spaces in urban areas. Complementing these findings, White et al. [120] analysed multi-country survey data on recreational visits to natural areas and mental health. Their results suggest protective effects of green and blue space exposure, with more frequent nature-based leisure trips associated with heightened wellbeing and lower distress.

In addition to mitigating distress, contact with green spaces may also confer positive mental health benefits mediated through social and behavioural pathways. Li et al. [121] demonstrated that residents' mental health significantly improved with higher green space exposure in a Chinese city, attributable to mechanisms including enhanced environmental perception, physical activity, and sense of community belongingness. Aligning with these conclusions, Qin

et al. [122] found positive green space-mental health linkages in Beijing neighbourhoods, mediated by physical activity and social cohesion. These findings illuminate potential mechanisms underlying observed nature wellbeing connections in urban contexts.

Equitable distribution and access to green spaces also emerge as a key variable influencing population mental health outcomes. Multiple studies highlight socioeconomic and spatial disparities in urban green space provisions [123–125]. Uneven access negatively impacts distressed communities and compounds mental health burdens linked to lower socioeconomic status. As urbanisation escalates globally, environmental justice perspectives warrant consideration in green space planning and resource allocation.

Recent research conducted worldwide resoundingly affirms the mental health significance of urban nature contact and green space access. Findings underline both protective and promotive pathways connecting green space exposure to enhanced psychological wellbeing. Moreover, equitable provision of quality green spaces emerges as an environmental justice imperative with implications for the population's mental health. As cities expand and densify, integrating nature-based solutions and ensuring access across socioeconomic gradients will be vital for nurturing healthy, sustainable communities. Looking ahead, further investigating specific mechanisms and at-risk populations can help refine nature-based mental health promotion policies and interventions in the urban context.

2.6.3 Air Quality and Public Health

The analysis of air pollution data and its myriad impacts on human health has garnered significant research attention amid rapid urbanisation and industrialisation worldwide. Both short- and long-term exposure to common air pollutants, including particulate matter (PM2.5, PM10), nitrogen dioxide (NO2), sulphur dioxide (SO2), and ozone (O3), have been linked to an array of acute and chronic health conditions [126, 127].

Multiple studies highlight the acute health effects associated with short-term exposure to these pollutants. Yang et al. [128] found that transient spikes in air pollution levels could precipitate respiratory diseases, stroke, and coronary artery disease, as well as exacerbate preexisting con-

ditions like asthma and chronic obstructive pulmonary disease (COPD). While Liu et al. [129] observed no significant immediate impact of short-term pollution exposure on subjective wellbeing, they noted substantially reduced subjective wellbeing following medium- and long-term exposures, suggesting potential delayed effects on mental health. Reviewing the interconnections between COVID-19 and air pollution, Katoto et al. [130] suggested that acute and chronic pollution exposure could influence COVID-19 transmission rates and severity by compromising lung function.

The impacts of sustained, long-term air pollution exposure appear even more significant. Cromar et al. [127] stressed that adhering to stringent air quality standards, such as those recommended by the American Thoracic Society, could mitigate substantial mortality and morbidity from long-term PM2.5 and ozone exposure. Similarly, Ali and Islam [131] reported that populations habitually exposed to PM2.5 and NO2 faced markedly higher COVID-19 infection and mortality rates due to chronically compromised lung function. Research quantifying effects across organ systems has reviewed physiological conditions exacerbated by chronic pollution exposure, including asthma, COPD, cardiovascular and metabolic diseases, alongside psychological outcomes like depression, stress, and dementia [126]. Studies have also elucidated geographically localised influences, such as the association between long-term air pollution exposure in Tehran, Iran, with increased blood pressure [132], or heightened mortality during high pollution periods in Bandung, Indonesia [133].

Cumulatively, the emerging research underscores the pressing need to improve air quality globally. Environmental regulations, public health policies, and awareness campaigns are needed to mitigate both acute and chronic health burdens imposed by ambient air pollution, particularly as industries expand worldwide. Timely interventions to curb pollution levels are imperative to safeguard public health amidst increasing urbanisation.

2.6.4 Climate Change and Seasonal Effects

The complex interrelationship between climate change, extreme weather events, seasonal variations, and human wellbeing is being increasingly scrutinised by researchers across disciplines.

Recent studies have investigated how phenomena exacerbated by climate change—including heatwaves, floods, storms, and shifting seasons—profoundly impact physical health, mental wellbeing, social structures, and economic stability.

Multiple analyses focused specifically on extreme weather events reveal clear links between these incidents and adverse health outcomes. For example, some models demonstrate that rapid temperature flux during heatwaves or storms can disrupt phenological cycles, compromising agricultural systems and ecosystems [134]. Assessments in Italy correlate volatile rainfall and snow melt with heightened risks of flooding and landslides [135]. Furthermore, evidence across Europe shows that increasing frequency and severity of extremely hot days leads to more heat-related illnesses [136].

In addition to overt threats to physical health, extreme weather events also take substantial tolls on mental wellbeing. Analysis of visitor data from the Toronto Zoo indicated that extreme weather events suppressed attendance, illustrating disruption to leisure activities tied to mental health [137]. Additionally, agricultural turmoil triggered by droughts and floods in Malawi increased economic vulnerability and poverty-related mental stress among rural households [134].

While extreme weather occasions acute distress, gradual seasonal shifts emerging from climate change also negatively impact wellbeing. A review of Ayurvedic traditions suggests that disruptions to seasonal cycles can undermine physical and mental balance, emphasizing that human health relies on aligning routines with predictable seasons.

Climate change and associated extreme weather pose grave interrelated threats to human wellbeing via physical health, mental welfare, social settings, and economic circumstances. Understanding and responding to these multifaceted impacts will become increasingly crucial as such events intensify worldwide. Integrative adaptation policies and practices are essential to promote resilience and safeguard human wellbeing in an era of climate change.

2.7 Challenges and Ethical Considerations

2.7.1 Data Privacy and Ethical Considerations

Beyond general data privacy, applying AI and machine learning to environmental and physiological data for wellbeing analysis, as explored in this research, brings forth additional ethical considerations. A key concern is algorithmic bias, which may arise from demographic imbalances in training data. If models are developed primarily on data from specific population groups, they risk producing inequitable or inaccurate predictions for underrepresented communities. This could compromise fairness in health assessments and exacerbate existing disparities.

Transparency and explainability are equally important when dealing with sensitive outcomes such as stress or mental health prediction. Incorporating explainable AI (XAI) techniques allows clinicians and users to interpret how decisions are made, enhancing trust and enabling better-informed actions. Accountability for model outputs is also critical: it must be clear who is responsible for the decisions or recommendations produced by AI systems, especially in health-related settings. This includes documenting model design, maintaining oversight mechanisms, and involving human reviewers when necessary.

Ethical deployment further requires secure data handling practices and robust informed consent protocols. Participants should be clearly informed about how their data will be collected, used, and stored, and be given opportunities to opt out without consequence. Ensuring fairness, promoting transparency, and maintaining accountability are thus central to responsible AI use in physiological signal analysis and wellbeing monitoring [138–144].

2.7.2 Data Collection and User Adherence

A limitation in much current research is the infrequent use of real-world data collection or testing [145]. Individuals experiencing artificially induced stress in controlled trials may not exhibit the same physiological patterns as those facing real-life wellbeing challenges. For real-world data collection, recruiting participants and motivating them to test new technologies and provide feedback can be difficult. Identifying users willing to try unfamiliar devices presents a

challenge [146]. User adherence and sustained engagement are crucial both for effective testing and for gathering sufficient data to train machine learning models, yet users may not immediately perceive the benefits. Engaging hard-to-reach communities, where mental wellbeing technologies could have significant impact, poses even greater recruitment and adherence difficulties. Device design plays a role; making devices small, portable, and aesthetically pleasing may promote engagement and allow use in various settings.

2.7.3 Interdisciplinary Collaboration

Challenges also exist on the diagnostic front. Affective sensing, the interpretation of emotional states from sensor data, is inherently complex due to subjectivity [147]. Sensor data alone may face difficulties in accurately interpreting individual emotional variances. While machine learning algorithms could potentially be tailored to individual users to accommodate these variances, this requires extensive data collection over time from each person, dependent on developing user-friendly data gathering methods. The effectiveness of mental wellbeing detection relies not only on the precision of classification algorithms but also crucially on the reliability of the sensors themselves. Inaccurate sensor readings can lead to erroneous model classifications. Nevertheless, integrating machine learning classifiers with commercially available sensors has shown promise, sometimes achieving stress detection accuracy comparable to clinical-grade sensors, highlighting AI's potential to enhance diagnostic accessibility [148].

2.7.4 Portability

Utilising sensors and feedback actuators within tangible interfaces designed for enhancing mental wellbeing presents several practical hurdles. A notable concern is potential device bulkiness arising from the need to accommodate sensors, a power source (battery), and feedback systems. While innovative feedback methods, such as Visio-Tactile feedback using actuated liquid metal droplets, are being explored for more dynamic and compact feedback, this technology is still developing and may not yet be readily integrable into current tangible devices [149].

2.7.5 Battery Life

Even if users are willing to utilise tools for mental wellbeing assessment, battery longevity remains a persistent challenge. Internet of Things (IoT) devices, intended to be compact and wearable, often have limited internal space due to essential microcontrollers and sensors. This leaves minimal room for larger batteries, necessitating frequent recharging. A potential workaround involves activating certain high-energy sensors only when triggered by specific preliminary actions, thus reducing continuous power drain but adding complexity to data collection. Until significant advancements occur in battery technology to extend operational life spans, the continuous collection of extensive behavioural or physiological data will remain constrained. Therefore, developing practical approaches to enhance power efficiency in device operation and data processing is essential.

2.8 Summary, Research Gaps, and Transition

This chapter synthesised contemporary research at the intersection of environmental factors and human wellbeing, emphasising the contributions of urban planning, environmental exposure, psychological frameworks, and digital technologies such as machine learning and the Internet of Things. Although there is growing enthusiasm for interdisciplinary and data-intensive approaches, the review uncovers persistent limitations in current methodologies. In particular, there is an absence of comprehensive models that cohesively integrate perspectives from urban design, environmental sciences, psychological research, and artificial intelligence. This gap restricts the ability to fully understand how environmental and psychosocial variables interact to influence health and wellbeing outcomes [138, 139]. Addressing these disciplinary boundaries is therefore essential to advance integrated monitoring and intervention strategies.

A further prominent issue concerns the type and resolution of data commonly used in environmental wellbeing research. Many existing studies depend on static, cross-sectional, or coarse-grained datasets, which are insufficient to track dynamic changes in environmental exposure or their impacts on individuals. Thus, there is a compelling need for longitudinal investigations that utilise real-time data streams from wearable technologies and environmental sensors, en-

abling finer-grained analysis of how environmental conditions influence physiological and psychological responses over time [141, 142]. Despite rapid progress in AI and machine learning, their full potential in this context remains largely unrealised. The application of high-capacity algorithms for large-scale, multimodal environmental health data remains in its infancy, particularly outside controlled research environments [143, 144].

Additionally, issues of equity and climate resilience are not adequately addressed in much of the current literature. Socioeconomic inequalities often determine access to beneficial environmental features such as parks, clean air, and sustainable infrastructure yet these factors are underrepresented in most studies. There is also a need for more emphasis on how urban planning can evolve to build resilience against climate-related disruptions, particularly in vulnerable populations. This doctoral research contributes to closing these gaps by employing a cross-disciplinary approach, analysing longitudinal, in-situ sensor data, and applying advanced machine learning techniques to model complex relationships between environmental exposure and wellbeing. Furthermore, it incorporates considerations of fairness and ethics in data collection and analysis, contributing to more responsible and inclusive digital health frameworks.

Having outlined the conceptual background and identified key research challenges, the next chapter presents the research methodology, detailing the experimental framework, data acquisition strategies, and analytical methods employed in this study.

Chapter 3

Classifying Mental Health Using Wearable Sensor Data

3.1 Introduction

Mental health disorders, particularly depression, remain a major public health concern. Prolonged stress is widely recognised as a risk factor for numerous physical and psychological conditions, including headaches, insomnia, and cardiovascular disease [89,150–152]. Although early and objective detection of affective disorders is critical, clinical assessment is still often based on subjective rating scales [153–155], highlighting the need for more robust, data-driven approaches (see Chapter 2, Section 2.2).

Wearable sensors now enable the collection of rich, continuous motor activity data, providing an objective window into individuals' behavioural patterns (see Section 2.5.2). Such data have been linked to depression and other affective disorders [156], but accurately interpreting these complex time series remains a challenge. Recent advances in machine learning have shown promise for revealing underlying patterns in motor activity and other physiological signals (see Section 2.5; Section 2.3.2) [157].

This chapter presents a comprehensive comparative study of machine learning and deep learning approaches for classifying depression based on motor activity data from the public Depresjon

dataset [156]. Extending previous work [158–161], this PhD research programme systematically evaluates a diverse set of classification algorithms, including LR, RF, Gradient Boosting GB, support vector machines SVM, TPOT AutoML, and the MLSTM-FCN Deep Learning DL model across multiple data representations (see Sections 2.5.4, 2.5.4.2, and 2.5.4.4). These include raw and normalised activity counts, statistical feature extraction, and dimensionality reduction using principal component analysis, with further segmentation into daytime and nighttime periods to capture potential diurnal effects (see Section 2.5.3).

By directly comparing the impact of different preprocessing strategies and classification models, this study aims to identify optimal methodological choices for distinguishing depressed patients from healthy controls in a wearable sensor context. Careful attention is given to addressing class imbalance and reporting a comprehensive set of evaluation metrics (see Section 2.8). The findings provide new insight into best practices for objective assessment of mood disorders, and inform subsequent chapters of this thesis.

3.2 Methodology

This section details the methodology employed for classifying individuals as depressed patients or healthy controls using motor activity data from the Depresjon dataset. It covers the dataset characteristics, the preprocessing steps undertaken to create different feature sets, the machine learning models applied, and the training and evaluation procedures.

3.2.1 Dataset and Preprocessing Procedure

Dataset Description. This study utilises the public Depresjon [1] dataset [156], chosen for its open-access nature and its use as a benchmark in related research [162, 163]. The dataset contains motor activity data collected using wrist-worn Actiwatch accelerometers (32 Hz sampling, recording movements greater than 0.05g) from 23 patients diagnosed with unipolar or bipolar depression (condition group) and 32 healthy controls. Activity counts, correlating with movement intensity, were recorded at one-minute intervals, with recording durations varying per par-

[1] Available online: https://datasets.simula.no/depresjon/, last accessed 20 May 2025.

ticipant (13-45 days). Demographic data (gender, age group) and clinical information (MADRS scores, inpatient/outpatient status, etc., in 'scores.csv') are also provided. Figures 3.1 and 3.2 illustrate the gender and age distributions within the dataset's condition and control groups, respectively.

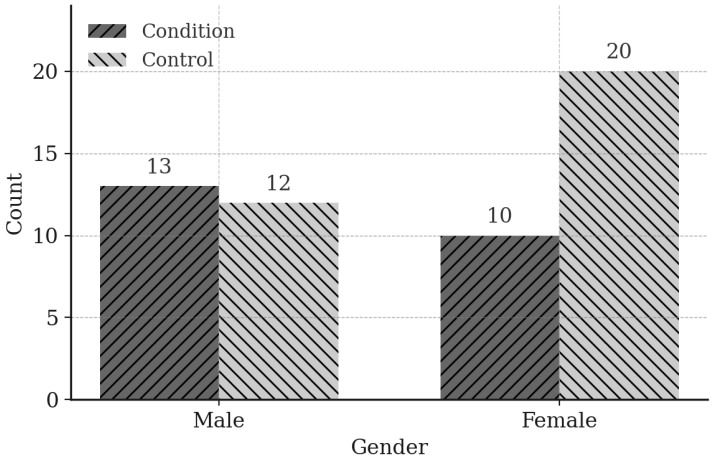


Figure 3.1: Bar Plot of Gender Distribution in Depressed vs. Healthy Control Groups

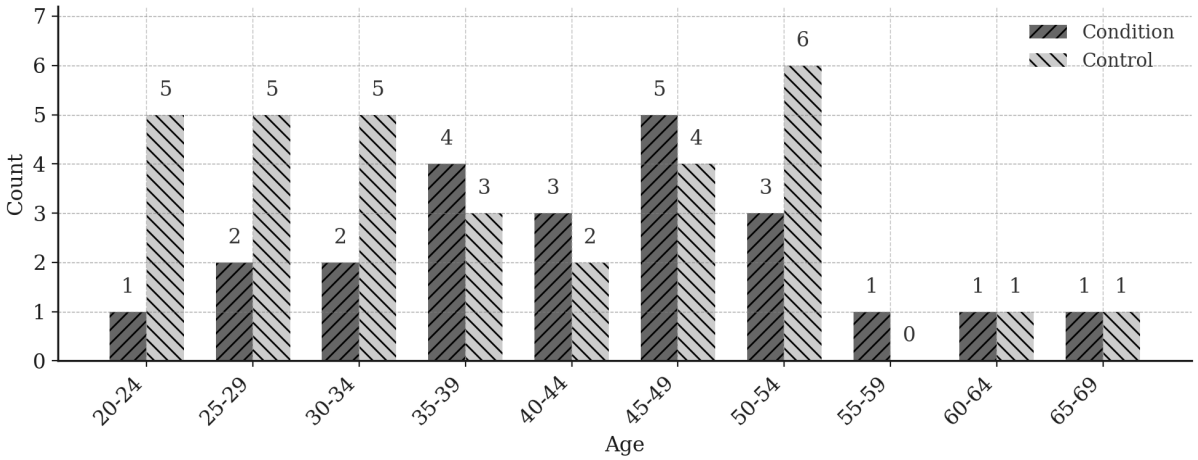


Figure 3.2: Age Group Distribution in Depressed and Healthy Control Groups

Data Segmentation and Initial Visualisation. To analyse potential diurnal variations, the raw data for each participant was segmented into 12-hour daytime (08:00-20:00) and nighttime (20:00-08:00) periods. This resulted in 1035 daytime samples and 1083 nighttime samples across all participants. Visual inspection of the average activity patterns (Figures 3.3 and 3.4) reveals that, while motor activity is broadly similar between depressed patients and healthy controls across both day and night, some group differences do emerge. Notably, healthy controls

tend to be more active than patients in the early morning hours (approximately 06:00 to 10:00). However, the substantial overlap in group averages throughout the rest of the day and night indicates that further, more sophisticated feature extraction and modelling are needed to robustly differentiate between groups.

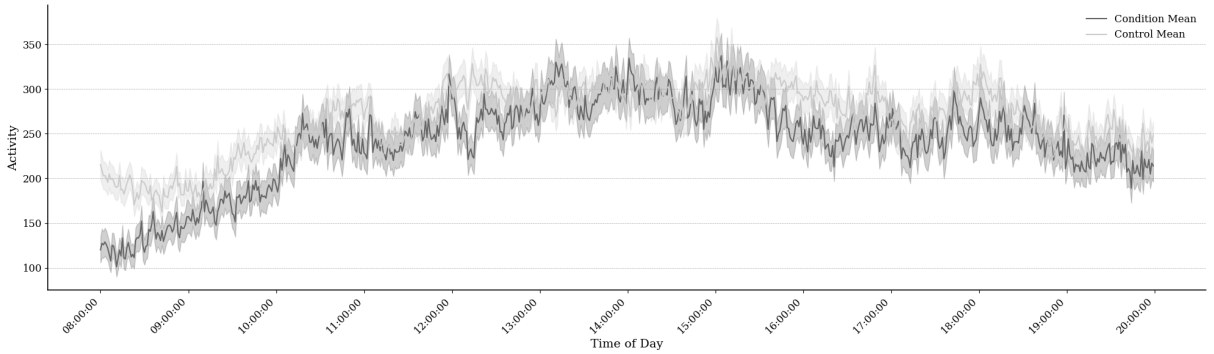


Figure 3.3: Daytime Activity Patterns of Depressed Patients vs. Healthy Controls

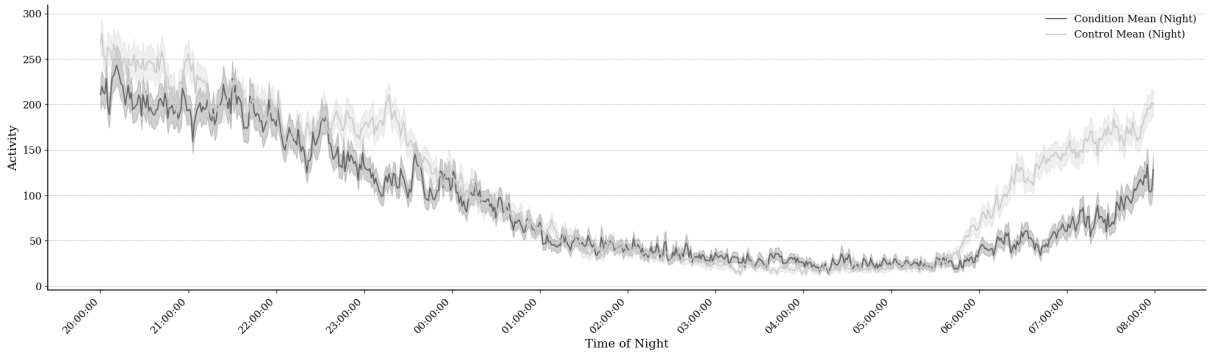


Figure 3.4: Nighttime Activity Patterns of Depressed Patients vs. Healthy Controls

Cross-correlation analysis, visualised as heatmaps (Figures 3.5 and 3.6), was also performed. The nighttime heatmaps suggested potentially stronger day-to-day consistency (higher correlations) in activity patterns for healthy controls compared to depressed participants, hinting at subtle differences despite the similarity in overall activity levels.

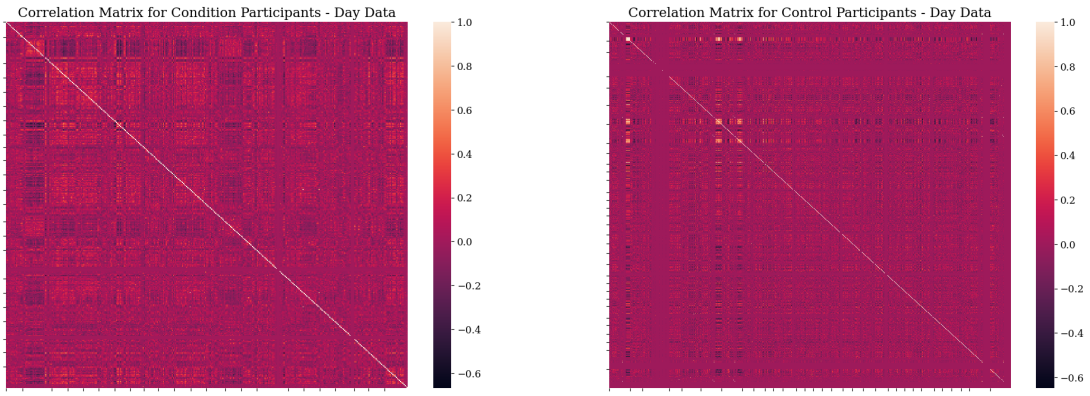


Figure 3.5: Heatmaps displaying daytime cross-correlation patterns for depressed patients (left) and healthy controls (right).

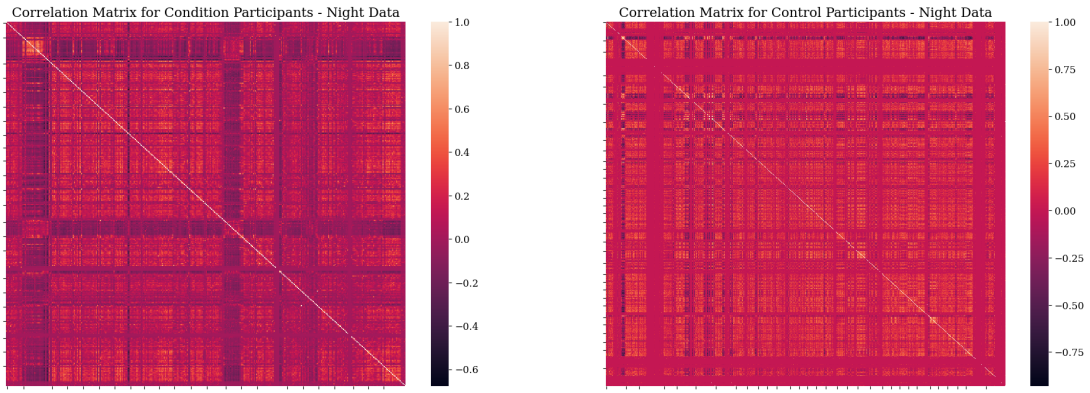


Figure 3.6: Nighttime cross-correlation heatmaps illustrating the relationships between daily activity patterns for depressed patients (left) and healthy controls (right).

Feature Set Creation. To enable a comprehensive comparison of preprocessing strategies, several distinct feature sets were derived from the segmented 12-hour activity periods for both daytime and nighttime. The most fundamental representation involved using the original raw activity counts for each minute, resulting in high-dimensional feature vectors of 720 features per sample that preserve the full temporal resolution of the data. Recognizing the potential for differences in measurement scale to bias model performance, these raw features were also normalized using MinMax scaling, transforming all activity values to the [0, 1] range, while maintaining the same dimensionality (720 features per sample).

Dimensionality reduction was explored by applying PCA [108, 111] to both the raw and normalized data; the number of retained components was determined by analysing the cumulative explained variance, with the top 300 principal components selected for the raw data to cap-

ture the majority of information (as shown in Figure 3.7). This yielded a feature vector of 300 principal components per sample.

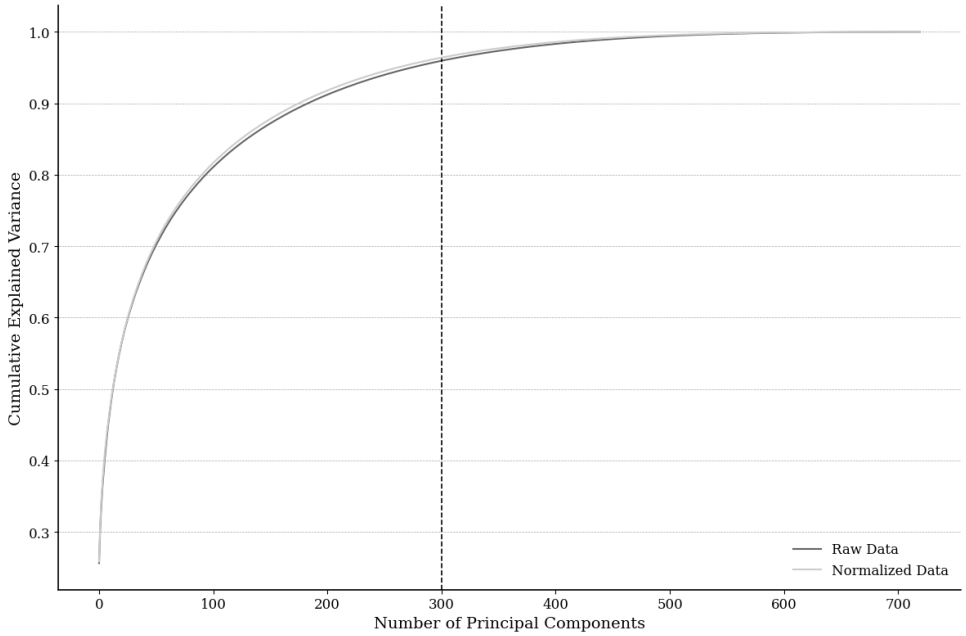


Figure 3.7: Cumulative Explained Variance Ratio for Principal Components of Raw Data and Normalized Raw Data (Daytime)

In addition to these direct representations, more abstract summaries of activity patterns were constructed by extracting statistical features from each hour within a 12-hour block. For each hour, seven statistical descriptors were calculated: minimum, maximum, mean, median, standard deviation, skewness, and kurtosis, yielding 84 features per sample (12 hours $\times$ 7 statistics). These hourly statistical features provide a more compact characterization of the data, emphasizing distributional properties and temporal variation [157]. As with the raw data, normalization was applied to the statistical features using MinMax scaling, ensuring comparability across features and participants while maintaining the dimensionality at 84 features per sample. To further reduce dimensionality and emphasize the most salient patterns, PCA was also performed on the normalized statistical features; the top 11 principal components were retained based on explained variance criteria (see Figure 3.8), resulting in a feature set of 11 components per sample.

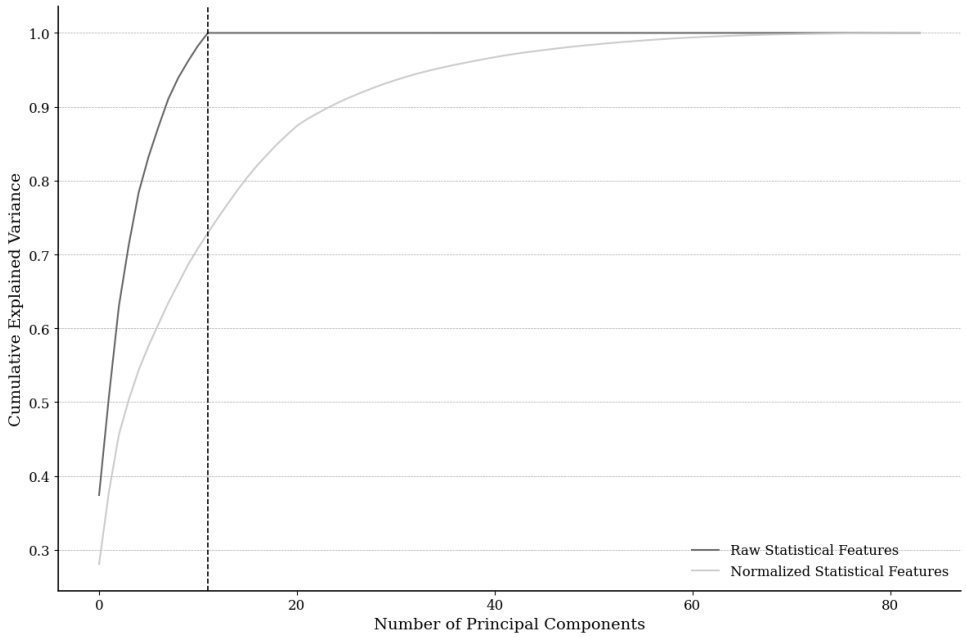


Figure 3.8: The Cumulative Explained Variance Ratio for the Principal Components of Statistical Features and Normalized Statistical Features (Daytime) .

By preparing these multiple feature sets, ranging from raw high-dimensional activity profiles to compact summaries of principal components of statistical features, this study systematically investigates how different representations of wearable sensor data influence the ability of machine learning and deep learning models to classify depression status. This layered approach to feature engineering aims to balance the retention of relevant information with the reduction of noise and computational complexity.

3.2.2 Proposed Models

A diverse set of machine learning models were selected for the task of classifying individuals as depressed patients or healthy controls based on their motor activity patterns, allowing for a comparative evaluation. The specific models employed in this study include classical approaches like Logistic Regression [105], Random Forest Classifier [49], Gradient Boosting Classifier [47], KNN, and Support Vector Machines (SVM) [164], as well as the automated machine learning tool TPOT [107], a Gaussian Process Classifier [48], and a deep learning model specifically designed for multivariate time series, the MLSTM-FCN [112]. The theoretical background for these algorithms is provided in Chapter 2 (Section 2.5).

The selection aimed to compare traditional, well-understood models with more recent deep learning architectures and automated approaches. TPOT, an AutoML system, automatically searches for optimal combinations of feature transformations, models, and hyperparameters using genetic algorithms, generating pipelines such as the one illustrated conceptually in Figure 3.9.

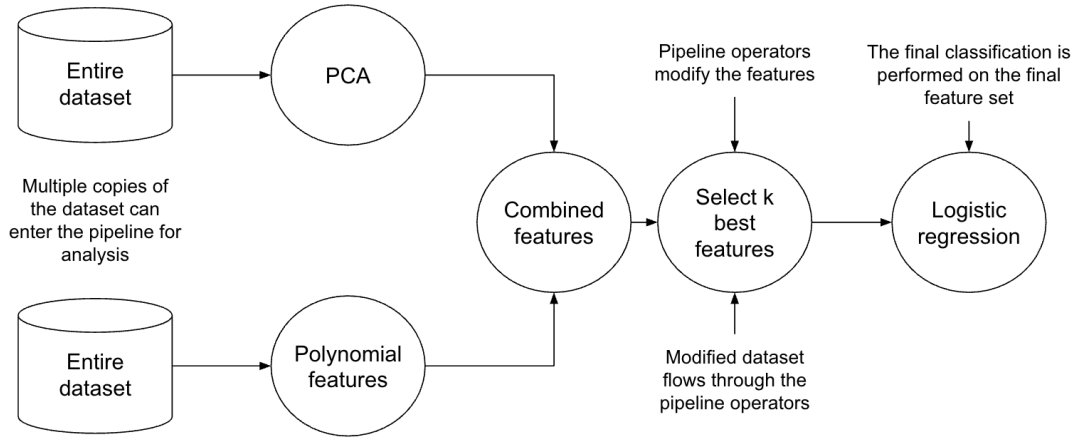


Figure 3.9: An illustrative representation of a tree-based pipeline generated by TPOT.

The MLSTM-FCN architecture [112], depicted in Figure 3.10, is a deep learning model developed for multivariate time series classification. In this study, the model was configured to accept input sequences representing 12-hour activity samples, either as raw minute-level activity counts or as transformed feature sets. The MLSTM-FCN combines two parallel branches: a LSTM network that learns temporal dependencies across the sequence, and a stack of 1D-CNN layers that extract salient local and hierarchical features from the same input. Outputs from both branches are concatenated and passed through fully connected layers to generate the final binary classification, distinguishing depressed patients from healthy controls.

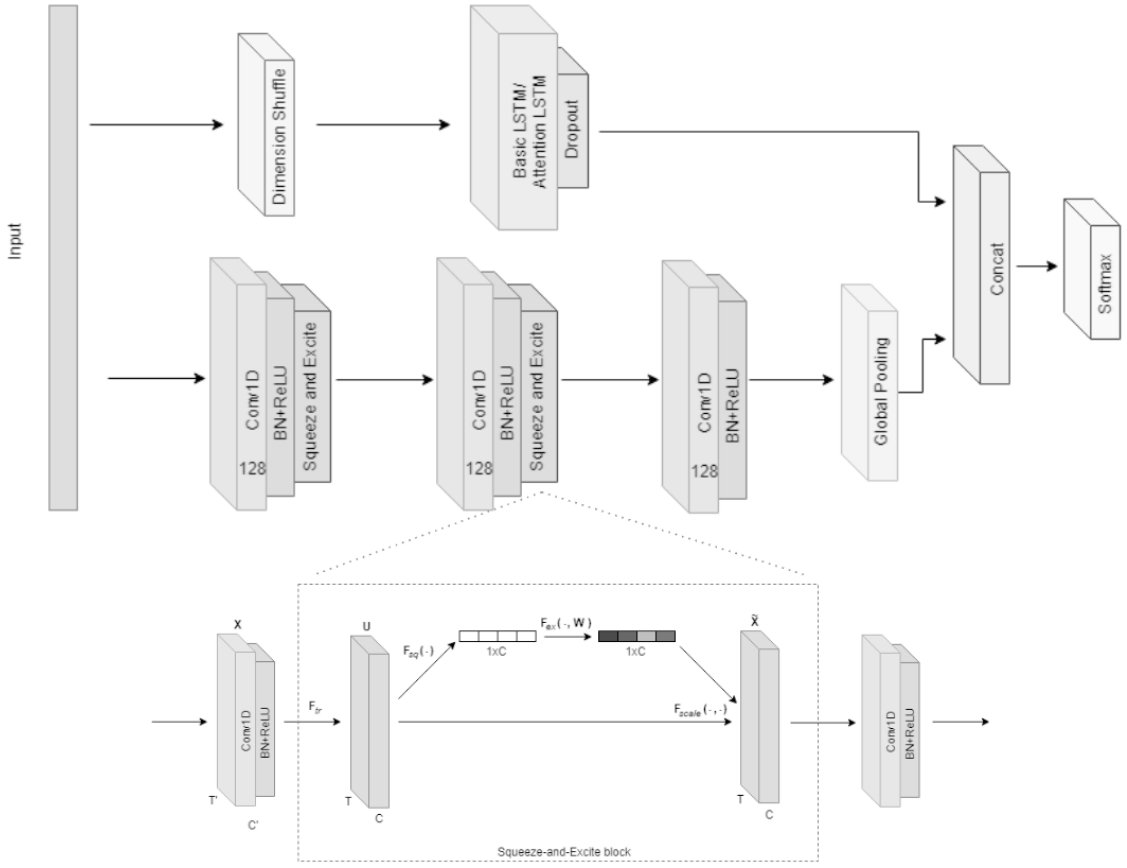


Figure 3.10: The MLSTM-FCN architecture.

3.2.3 Training and Testing Steps

Following model selection and data preprocessing, a structured approach was adopted for training, validating, and testing the classifiers. The preprocessed datasets were split into training and test sets to ensure robust evaluation. A 5-fold cross-validation technique was implemented during the training phase, particularly for hyperparameter tuning using GridSearchCV technique [165]. The specific hyperparameter grids explored, along with the total number of parameter combinations and the optimal configurations selected for each algorithm, are summarised in Table 3.2.

The performance of the trained models was evaluated using standard classification metrics: Accuracy, Precision, Recall, and F1-score. These metrics provide a comprehensive assessment of the model's ability to correctly classify both the condition and control groups, considering aspects like class imbalance. To potentially enhance classification accuracy, threshold selection

based on the Receiver Operating Characteristic (ROC) curve analysis was implemented after initial model training. Additionally, ensemble predictions using a Vote Classifier, combining outputs from multiple individual models, were explored [165].

3.3 Results

This section details the stepwise evaluation of machine learning and deep learning models for classifying depression using motor activity data from the Depresjon dataset. The analysis follows a structured approach: initial cross-validation to select the best feature representation, model fine-tuning, training and evaluation analysis, and comprehensive performance reporting including threshold optimization.

Cross-Validation to Select Optimal Feature Set. The first phase involved evaluating all model-feature set combinations using 5-fold cross-validation. As summarised in Table 3.1, the "Normalised Statistical Features" representation—particularly for nighttime data—emerged as the most effective across multiple model types, consistently yielding the highest cross-validation accuracies. This finding motivated the use of this feature set for all subsequent model comparisons and hyperparameter optimization.

Table 3.1: Average 5-Fold Cross-Validation Accuracies (%) for Daytime/Nighttime datasets across different feature sets. (RD: Raw Data, NRD: Normalised Raw, PCA-TRD: PCA Raw, SD: Statistical Data, NSD: Normalised Statistical, PCA-TSD: PCA Statistical).

Model	RD	NRD	PCA-TRD	SD	NSD	PCA-TSD
LR	59.32/60.94	60.58/60.48	56.71/60.29	65.31/67.22	**69.66**↑/**75.99**↑	**49.95**↓/**56.14**↓
RF	75.17/75.34	75.17/75.25	**68.60**↓/**73.03**↓	75.07/**78.29**↑	**75.75**↑/77.93	69.28/76.45
GB	71.49/**70.08**↓	71.40/70.08	**69.66**↓/72.67	75.07/**78.39**↑	**75.46**↑/78.21	70.24/75.62
KNN	**67.92**↓/69.90	68.70/**68.98**↓	68.31/70.18	69.86/73.69	**71.01**↑/**74.70**↑	68.21/73.96
SVM	72.27/71.92	71.69/**70.26**↓	72.37/72.11	65.80/70.35	**72.95**↑/**78.02**↑	**65.22**↓/71.37
GP	65.02/64.81	**72.65**↑/67.77	65.31/64.90	64.83/64.17	72.46/**77.38**↑	**64.44**↓/**64.44**↓
Hybrid Model	74.78/72.66	74.01/**72.11**↓	71.88/74.42	73.43/77.09	**75.07**↑/**77.19**↑	**70.14**↓/76.63
TPOT Model	**68.21**↓/**70.55**↓	68.79/71.28	72.66/72.94	73.43/76.82	**73.62**↑/**77.28**↑	69.76/75.16
MLSTM-FCN	70.92/74.79	70.82/73.77	**66.76**↓/**72.94**↓	68.41/**76.09**↑	**73.24**↑/75.90	N/A

Model Fine-Tuning and Training Analysis. After identifying the best-performing feature set, hyperparameter optimization was conducted for the classical machine learning models (RF, GB, KNN, SVC) using GridSearchCV on the training set. The hyperparameter configurations are provided in Table 3.2. The MLSTM-FCN deep learning model, by contrast, was not fine-tuned

using grid search due to the computational cost and standard practice of training such models with fixed hyperparameters. Instead, its training process is illustrated in Figure 3.11, which plots training and validation accuracy and loss over epochs.

Table 3.2: Summary of Hyperparameter Exploration and Optimal Configurations for Machine Learning Models

Algorithm	**Parameters Explored**	**Total Parameter Sets**	**Best Parameters**
RF	– bootstrap: [True, False] – max depth: [None, 3, 8] – max features: [”auto”, ”log2”,”sqrt”] – min samples split: [2, 5,10] – min samples leaf: [1, 2, 4] – criterion: [”gini”, ”entropy”] – n estimators: [10, 50, 100, 200]	1728	– bootstrap: True – criterion: ’entropy’ – max depth: None – max features:’auto’ – min samples leaf: 1 – min samples split: 2 – n estimators: 100
GB	– learning rate: [0.01, 0.1, 0.2] – min samples split: [2, 5, 10] – min samples leaf: [1, 2, 4] – max depth: [3, 5] – max features:[’log2’,’sqrt’, ’auto’] – n estimators: [50, 100, 200]	486	– learning rate: 0.1 – max depth: 3 – max features: ’log2’ – min samples leaf: 4 – min samples split: 10 – n estimators: 100
KNN	– n neighbors: [1, 2, . . . 49]	49	– n neighbors: 25
SVM	– C: [0.1, 1, 10, 100] – kernel: [’linear’, ’rbf’, ’sigmoid’] – gamma:[’scale’, ’auto’] – coef0:[0, 1, 2, 3] – shrinking:[True, False]	192	– C: 1 – kernel: ’rbf’ – gamma: ’scale’ – coef0: 0 – shrinking: True

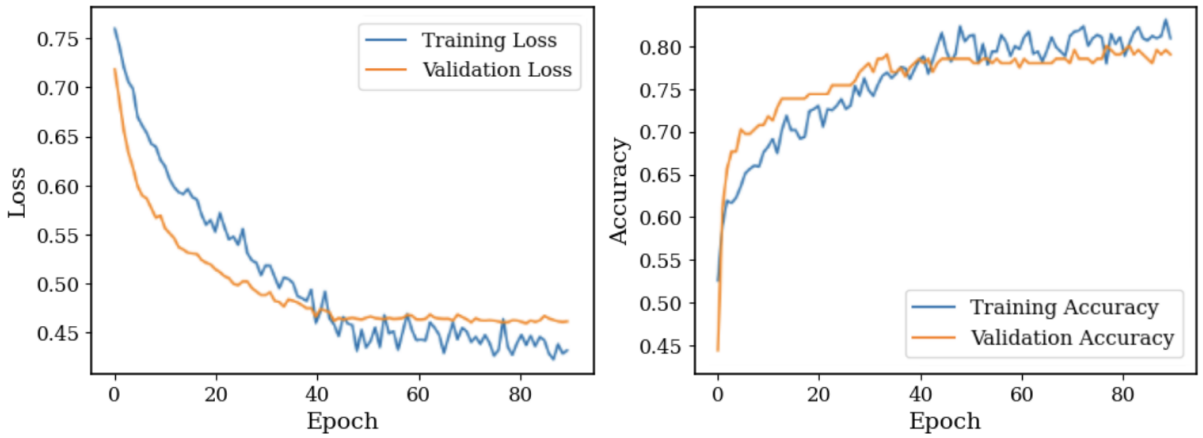


Figure 3.11: Training Curves for the MLSTM-FCN Model Over Epochs

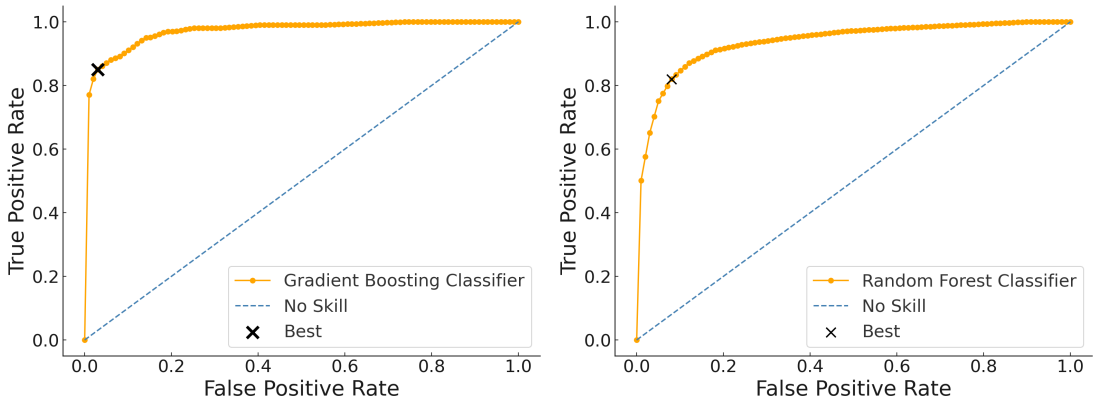
Train and Test Accuracy Comparison. Once optimal hyperparameters were selected for each classical model (and the MLSTM-FCN trained to convergence), all models were evaluated on the held-out test set using the "Normalised Statistical Features" (nighttime) data. Table 3.3 presents train and test accuracies for each model. The Random Forest and Gradient Boosting classifiers outperformed other models, with test accuracies of 83.41% and 82.03%, respectively, while MLSTM-FCN achieved 77.88%.

Table 3.3: Train and Test Accuracies (%) for Fine-Tuned Models on Normalised Statistical Features Nighttime Dataset.

Model	Train Accuracy	Test Accuracy
Random Forest	**99.42**↑	**83.41**↑
Gradient Boosting	92.61	82.03
KNN	79.45	**73.27**↓
SVM	84.53	80.18
TPOT Model	85.79	74.65
MLSTM-FCN Model	**84.18**↓	77.88

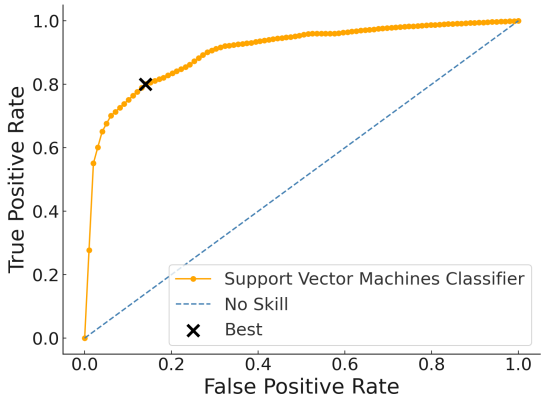
Threshold Optimization and Comprehensive Metrics. To further refine model evaluation, ROC curves were generated for each optimized model, enabling threshold selection that balances sensitivity and specificity (Figure 3.12). Thresholds were chosen to maximize the geometric mean of the true positive rate (TPR) and (1 - false positive rate). Using these opti-

mal thresholds, final performance metrics (Precision, Recall, and F1-score) for both "Control" and "Condition" groups are summarised in Table 3.4. This comprehensive view confirms the robustness of RF and GB classifiers, and highlights specific strengths of the MLSTM-FCN, particularly in "Condition" recall.

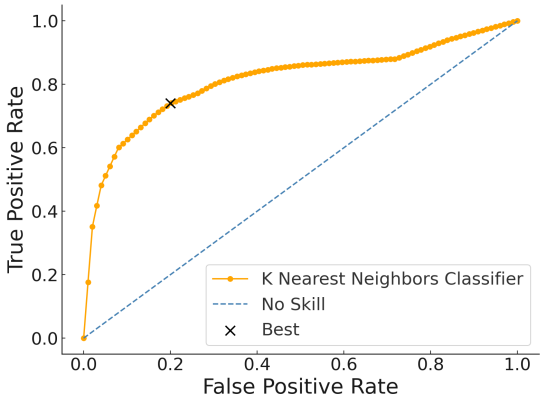


(a) Gradient Boosting Classifier

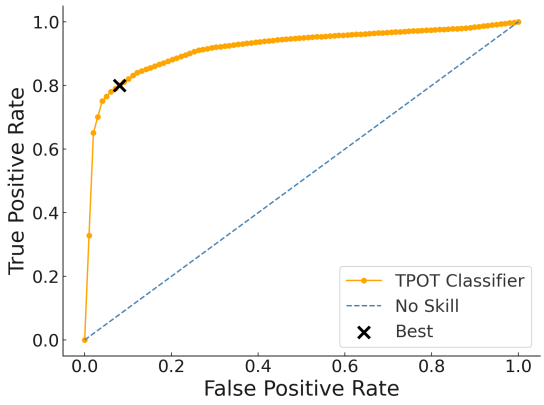
(b) Random Forest Classifier



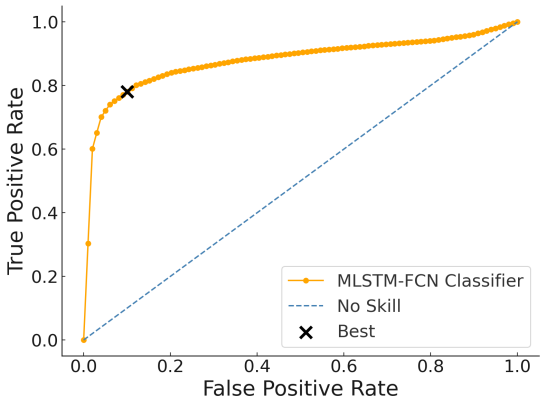
(c) SVM Classifier



(d) KNN Classifier



(e) TPOT Classifier



(f) MLSTM-FCN

Figure 3.12: ROC curves for the optimized machine learning models on the test set.

This stepwise analysis confirms that careful feature selection, hyperparameter tuning, and threshold optimization significantly enhance model performance for depression classification

Table 3.4: Summarises the Precision, Recall, and F1-score metrics for both the training and test datasets across various machine learning models (using optimal thresholds).

Model	Group	Train			Test		
		Precision	Recall	F1-score	Precision	Recall	F1-score
Random Forest Classifier	Control	**99%**↑	**100%**↑	**100%**↑	85%	**93%**↑	**89%**↑
Random Forest Classifier	Condition	**100%**↑	**98%**↑	**99%**↑	**79%**↑	62%	69%
Gradient Boosting Classifier	Control	96%	93%	94%	**90%**↑	83%	87%
Gradient Boosting Classifier	Condition	88%	93%	90%	68%	79%	**73%**↑
K Neighbors Classifier	Control	**86%**↓	**75%**↓	**80%**↓	89%	**67%**↓	**77%**↓
K Neighbors Classifier	Condition	**65%**↓	78%	**71%**↓	**52%**↓	**82%**↑	**64%**↓
Support Vector Machines	Control	92%	84%	88%	88%	74%	80%
Support Vector Machines	Condition	76%	87%	81%	56%	76%	65%
TPOT Classifier	Control	89%	89%	89%	80%	81%	81%
TPOT Classifier	Condition	85%	**77%**↓	81%	57%	**68%**↓	62%
MLSTM-FCN Classifier	Control	91%	83%	87%	87%	80%	83%
MLSTM-FCN Classifier	Condition	74%	86%	80%	62%	73%	67%

using wearable motor activity data. The strongest performance was observed for classical ensemble models (RF, GB), while the deep learning model provided competitive results, especially in recall for the "Condition" group. These findings reinforce the value of systematic comparative evaluation in health informatics applications.

The subsequent chapter introduces the development and characteristics of the "EnviroWellBeing" dataset, created for the environmental wellbeing analyses presented later in this thesis.

Chapter 4

EnviroWellBeing Dataset

4.1 Introduction

Recent advancements in sensor technology have enabled unprecedented opportunities for monitoring environmental and physiological factors relevant to human health (see Chapter 2). However, current datasets still fall short when it comes to comprehensively capturing the immediate and nuanced effects of specific environmental exposures on individuals using real-time, multimodal sensor data [166, 167].

A major limitation of existing resources is their lack of either sufficient granularity, breadth of measurements, participant diversity, or contextual richness necessary to address the complex research questions posed in this thesis. For instance, the DigitalExposome dataset [168] integrates environmental and physiological signals, but its final sample size after exclusions is small (12 participants), and its coverage of both physiological and contextual variables is limited relative to the requirements for robust machine learning analyses. Similarly, the WESAD dataset [89] is widely used for affective computing but focuses primarily on laboratory conditions with just 15 participants, and lacks simultaneous real-world environmental data. Other datasets, such as Urban Wellbeing [169] or The City [170], may offer real-world context, but are constrained by the absence of high-resolution physiological data or are limited to brief, static observations.

Given these limitations, there remains a significant gap in resources suitable for the integrated,

high-dimensional analysis required to explore how environmental factors shape human wellbeing in everyday life. To address this gap, this research undertook the development of the **EnviroWellBeing Dataset** [1] (RQ1): a new, comprehensive dataset specifically designed to overcome the shortcomings of prior work. The EnviroWellBeing Dataset includes synchronised, high-frequency recordings of both environmental and physiological parameters, collected from a diverse group of 53 participants in both real-world and controlled settings, with carefully designed experimental protocols to isolate the effects of environmental exposures. This experiment took place under the ethical approval number 21/22-53, please see Appendix B.4 for further details.

This chapter provides an overview of the EnviroWellBeing Dataset, beginning with the rationale and motivation for its creation, and then detailing the experimental design, equipment and sensor technologies used (see Section 4.2), participant recruitment and data collection procedures (Section 4.3), and initial data quality assessments (Section 4.5). By introducing this resource, the chapter sets the foundation for the subsequent application of machine learning and deep learning models to analyse the impact of environment on wellbeing.

4.2 Equipment

Participants were monitored using an integrated suite of wearable sensors and an environmental data logger, allowing for high-resolution collection of physiological and environmental data throughout the experiments. Figure 4.1 presents the equipment configuration used in the study.

[1] Available Online: EnviroWellBeing Dataset, last accessed 20 May 2025

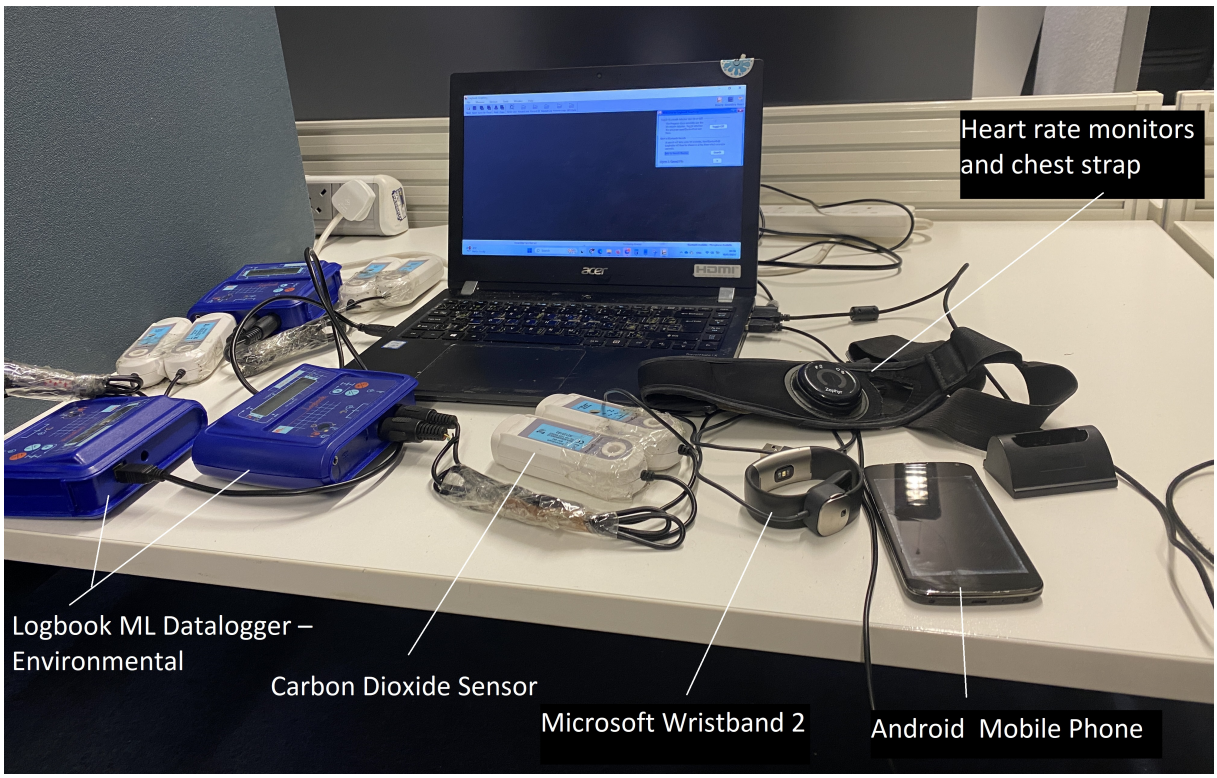


Figure 4.1: Sensor equipment setup for data collection post-experiment. Devices are labelled: (A) BioHarness chest belt sensor; (B) Microsoft Band 2 wristband sensor; (C) Logbook ML+ environmental data logger. This arrangement enabled comprehensive, synchronised monitoring of physiological and environmental parameters.

4.2.1 ChestBelt Sensors

Each participant was fitted with a BioHarness[2] chest belt sensor (Figure 4.2), an advanced wearable capable of simultaneously measuring a broad range of physiological parameters at high resolution. The device provided real-time acquisition of heart rate (HR) and breathing rate (BR) at a primary sampling rate of 1 Hz, supplemented by high-fidelity ECG waveforms sampled at 250 Hz, and respiratory rate data at 18 Hz. In addition to these core metrics, the BioHarness continuously recorded skin temperature, body posture, general activity level, and peak acceleration, affording a multi-dimensional perspective on the participant's physical state.

Notably, the sensor's ECG functionality enabled the extraction of R-to-R intervals—a critical index for heart rate variability (HRV) analysis and for evaluating autonomic and cardiovascular responses to environmental exposures. The combination of postural and activity data, together

[2]Available online: BioHarness 3.0, last accessed 20 May 2025

with ECG and temperature measurements, allowed for nuanced differentiation between baseline physiological states and changes triggered by environmental or experimental conditions. The ergonomic chest-worn form factor ensured reliable contact and signal integrity, supporting the collection of uninterrupted, synchronised data streams throughout the experimental sessions.

Figure 4.2: BioHarness chest belt sensor used for continuous multi-parameter physiological monitoring.

4.2.2 Wristband Sensors

To augment physiological and environmental monitoring, the Microsoft Band 2[3] wristband sensor (Figure 4.3) was deployed on each participant. This second-generation fitness tracker, equipped with multiple sensors, enabled the continuous collection of both bodily and ambient variables. Data streams from the wristband included electrodermal activity (EDA) for assessment of sympathetic nervous system activity, heart rate, and skin temperature, all sampled at 8 Hz to capture subtle fluctuations in physiological state. The device also measured ultraviolet (UV) exposure and ambient environmental noise, offering valuable context regarding light and sound environments to which participants were exposed.

A built-in three-axis accelerometer facilitated detailed monitoring of movement, posture transitions, and overall physical activity patterns. Air pressure readings provided additional en-

[3] Available online: Microsoft Band 2, last accessed 20 May 2025

vironmental context, such as changes in altitude or weather conditions during the study. The wristband's Bluetooth-enabled synchronisation with mobile devices allowed for secure data storage and real-time review, ensuring that no data were lost during collection. This rich dataset contributed to a multi-faceted understanding of participant experience within each experimental scenario.

Figure 4.3: Microsoft Band 2 wristband sensor used for physiological, environmental, and movement monitoring.

4.2.3 Data Loggers

Environmental conditions were continuously monitored using the Logbook ML+ data logger,[4] a robust device operating at a sampling rate of 0.5 Hz (Figure 4.4). This logger enabled the high-precision measurement of ambient temperature, light intensity, and sound levels, essential for contextualising the physiological responses observed in participants. Furthermore, the device recorded concentrations of key air pollutants, including volatile organic compounds (VOCs), carbon monoxide (CO), carbon dioxide (CO_2), nitric oxide (NO), nitrogen dioxide (NO_2), and sulphur dioxide (SO_2). The inclusion of these variables was critical for assessing the air quality experienced by each participant during both outdoor (in situ) and indoor (in vitro) experimental sessions.

Additional features of the data logger included auto-recognition of all attached sensors and flexible recording modes such as autoroll and snapshot permitting detailed tracking of transient environmental events. The capacity for fast recording and buffering allowed for the capture of

[4]Available online: Logbook ML Datalogger – Basic, last accessed 20 May 2025

rapid changes, enhancing the temporal resolution of environmental data during the experiments.

Figure 4.4: Logbook ML+ environmental data logger used for continuous measurement of ambient conditions and air quality indicators.

Taken together, this comprehensive instrumentation enabled the creation of a synchronised, multi-modal dataset, capturing the dynamic interplay between environmental exposures and physiological responses. The advanced features and high-resolution capabilities of each device ensured the robustness of data required for in-depth analysis of environmental health dynamics.

4.3 Data Collection

The EnviroWellBeing dataset was compiled to investigate the short-term impact of environmental conditions on human wellbeing, following established research highlighting the influence of urban pollutants, acoustic disturbances, and other environmental stressors on human health. The study collected data from 53 participants through a series of in situ (outdoor) and in vitro (indoor) experiments, each designed to systematically capture physiological and environmental responses using advanced wearable and environmental sensors (see Section 4.2).

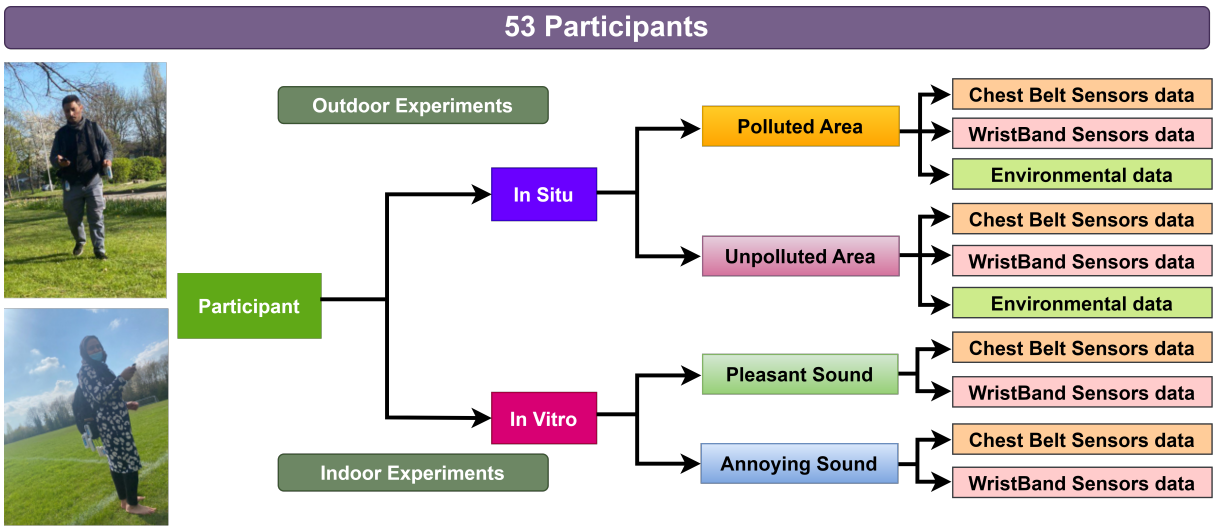


Figure 4.5: Overview of the EnviroWellBeing dataset. The figure depicts the integration of wearable physiological sensors and environmental data loggers deployed across a sequence of in situ (outdoor) and in vitro (indoor) experiments.

In Situ Experiments. Each participant completed two outdoor walking trials, traversing both a polluted urban environment and a green, unpolluted space, within the same geographical area. The polluted route was selected for its higher expected levels of air pollutants, noise, and urban stressors, while the green route followed paths through parkland or quieter residential areas, serving as a naturalistic baseline. This design is consistent with prior approaches to studying the environmental determinants of physiological and psychological stress. Figure 4.6 illustrates the exact paths followed in each condition.

Note: The GPS sensor was non-functional for the majority of the experiment, leading to the exclusion of GPS data from our dataset.

Figure 4.6: Outdoor experiments: (Left) Green (unpolluted) route; (Right) Urban (polluted) route. Maps show the designated walking paths for controlled comparison.

Throughout the outdoor sessions, participants wore chest belt sensors, wristband sensors, and

carried an environmental data logger (Section 4.2). The protocol initially included GPS tracking, but technical limitations led to its exclusion from the final dataset, a challenge also acknowledged in related environmental sensor studies [166]. Despite this, the combination of synchronised physiological and environmental data streams provided a comprehensive record of exposures and responses.

In Vitro Experiments. The indoor experiments focused on controlled auditory exposures. Using Adobe Audition software, custom two-minute audio clips were created to simulate urban (polluted) and green (unpolluted) sound environments. The polluted environment clip included disruptive noises such as car horns, engine and traffic sounds, and emergency sirens, while the green environment clip comprised tranquil sounds like birdsong, rustling leaves, and flowing water. The use of environmental sound simulation in laboratory settings for physiological research is consistent with previous studies [171,172]. The full list of auditory stimuli is provided in Table 4.1, as requested by examiners.

Table 4.1: Summary of Auditory Stimuli Used in In Vitro Experiments

Condition	Specific Sounds Included
Annoying	Car horns, engine noise, ambulance sirens, police sirens, construction sounds
Pleasant	Birdsong, wind in trees, running water, rustling leaves, distant soft voices

Each participant was exposed to both auditory environments in randomised order, with physiological signals such as heart rate and electrodermal activity continuously monitored throughout. The standardised 2-minute exposure per condition ensured consistency across all participants, facilitating within-subject comparisons.

Experiment Duration and Considerations. The duration of the outdoor experiments varied depending on each participant's walking pace, resulting in a wide range of exposure times, especially in the polluted environment. As shown in Figure 4.7, the box plot summarises the distribution of durations for each experimental condition (polluted, unpolluted, annoying, and pleasant). Participants generally spent longer and more variable times in the polluted area (median around 6–7 minutes, with several outliers), while durations in the unpolluted area were

shorter and more consistent (median approximately 2–3 minutes). In contrast, the indoor (in vitro) auditory exposure experiments had fixed durations of 2 minutes per condition for all participants, as indicated by the minimal spread in the box plots for 'Annoying' and 'Pleasant' conditions. Any technical issues encountered during data collection—such as sensor malfunctions due to battery depletion—were carefully documented and resolved during data preprocessing (see Section 4.4).

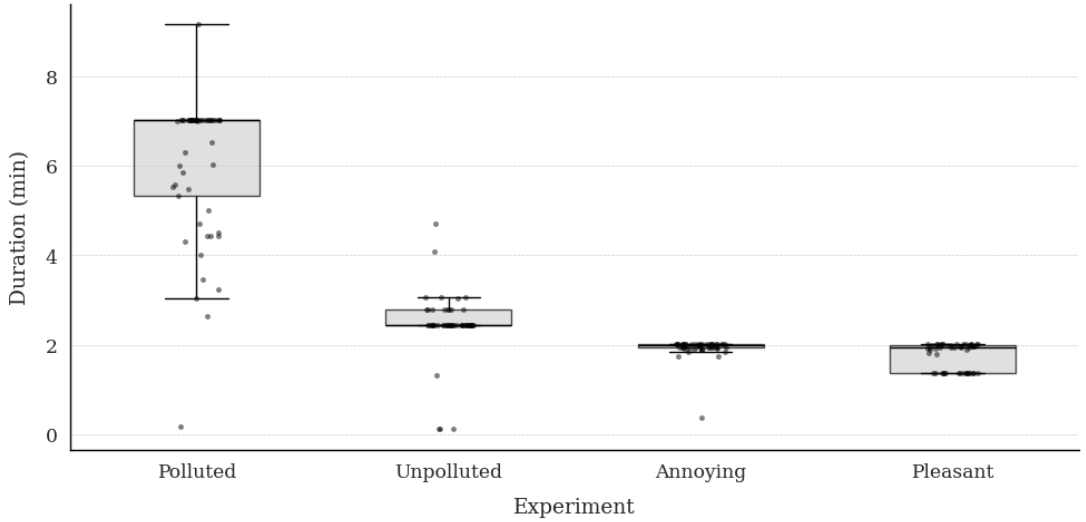


Figure 4.7: Distribution of participant durations by experimental condition. Each point reflects the time a participant spent in each environment (polluted, unpolluted, or auditory exposure), capturing natural variability.

By integrating synchronised physiological and environmental measurements in both natural and controlled conditions, this dataset enables a detailed analysis of how environmental exposures influence both subjective wellbeing and objective physiological responses, thereby directly addressing Research Question 1 (RQ1). The approach, which systematically controls for environment, stimulus, and individual pacing, provides a strong foundation for the machine learning analyses presented in subsequent chapters.

4.4 Data Preprocessing

Following data collection, extensive preprocessing was required to prepare the EnviroWellBeing dataset for analysis. The aim was to harmonise and clean multimodal time-series data from multiple sensors to ensure robust analysis of environmental impacts on wellbeing. The over-

all sequence of preprocessing steps, including feature extraction, harmonisation, cleaning, and imputation, is summarised in the data preprocessing flowchart presented in Figure 4.8.

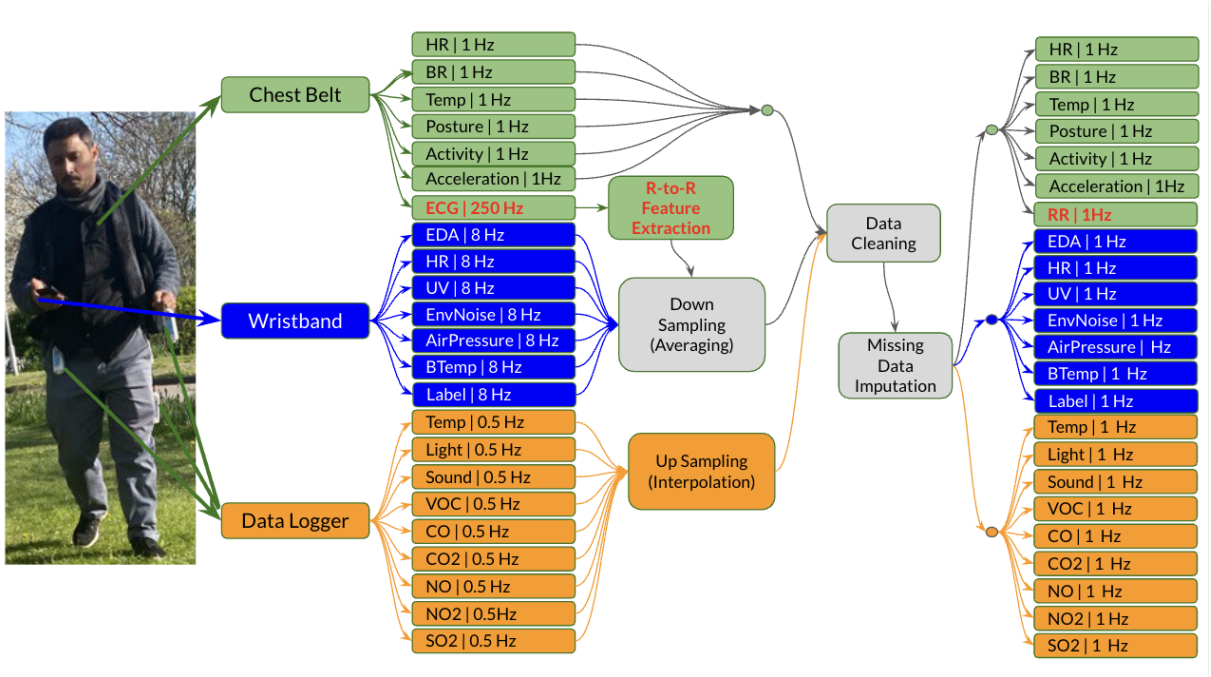


Figure 4.8: Data preprocessing flowchart. This pipeline summarises the harmonisation, cleaning, and feature extraction processes applied to synchronise and prepare multimodal sensor data for analysis.

The initial step was to select the most relevant physiological and environmental features from the comprehensive sensor data. The chest belt provided HR, breathing rate (BR), skin temperature, posture, activity, acceleration, and R-to-R intervals derived from ECG. The wristband sensor contributed EDA, ultraviolet (UV) exposure, environmental noise, air pressure, body temperature, and self-reported stress level labels. Environmental parameters captured by the data logger included ambient temperature, light level, sound level, and concentrations of gases such as CO, CO_2, NO, NO_2, SO_2, and VOCs.

Accurate analysis of heart rate variability (HRV) required extracting R-to-R intervals from raw ECG signals. As illustrated in Figure 4.9, this process involved identifying R-peaks in the ECG waveform, calculating the intervals between successive peaks, using histogram analysis to flag outliers, and then applying backward filling to address gaps or aberrant values. R-to-R interval values were subsequently averaged within each one-second interval to align with other synchronised features. The BioSPPy Python library was used to facilitate these biosignal

processing steps.

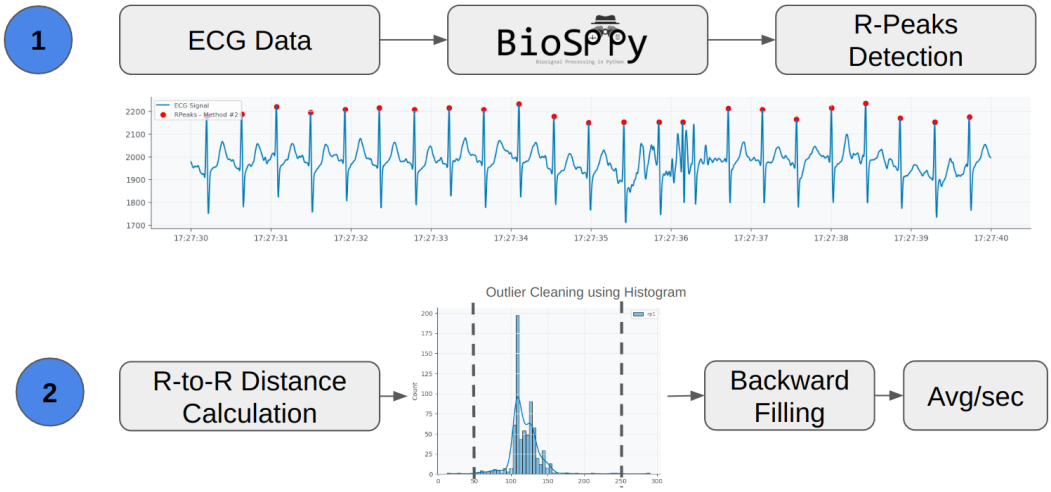


Figure 4.9: Extracting R-to-R features from ECG data: workflow for detecting R-peaks, computing R-to-R intervals, removing outliers, and resampling to 1 Hz.

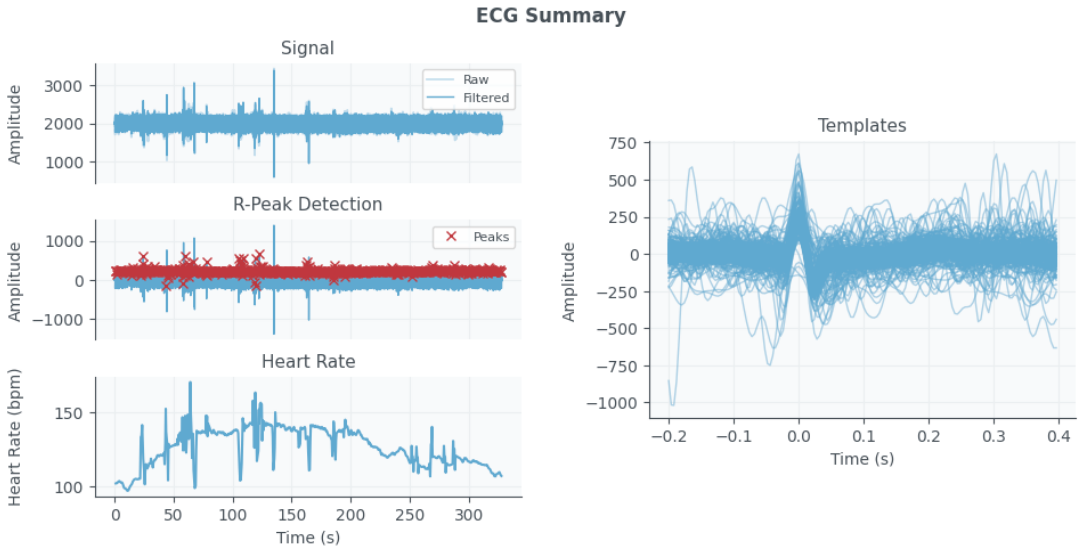


Figure 4.10: Summary output from BioSPPy: Raw ECG, detected R-peaks, derived heart rate, and average QRS template. This illustrates the ECG feature extraction and quality control process.

To enable synchronised, second-by-second time-series analysis across all sensor streams, all features were resampled to a unified 1 Hz frequency. This involved downsampling high-frequency signals, such as wristband features originally sampled at 8 Hz, by averaging values within each one-second interval, while categorical labels were rounded as needed. For low-frequency data, such as the environmental logger sampled at 0.5 Hz, linear interpolation was

employed to estimate intermediate values, creating a continuous 1 Hz signal that aligns with the physiological data.

All selected features then underwent careful cleaning to identify and address noise, artefacts, and outliers. Visual inspections using line plots, box plots, and histograms were performed to detect physiologically implausible values or abnormal patterns. For example, heart rate data were assessed via plots for sudden spikes or drops and distributions were compared across participants and conditions. Outliers were handled by applying defined physiological thresholds, with affected data points set to missing if they exceeded these bounds. The data cleaning process is exemplified in Figures 4.11–4.13, which demonstrate the visualisation, detection, and removal of outliers in heart rate data. The impact of cleaning is further illustrated by comparing histograms and KDE plots before and after correction.

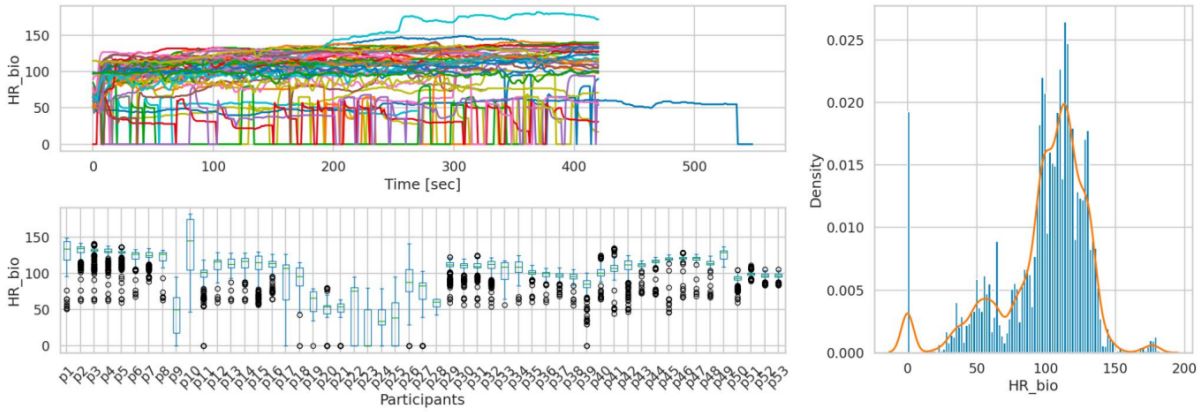


Figure 4.11: Line, box, and histogram plots used to visually inspect and detect outliers in heart rate data during cleaning.

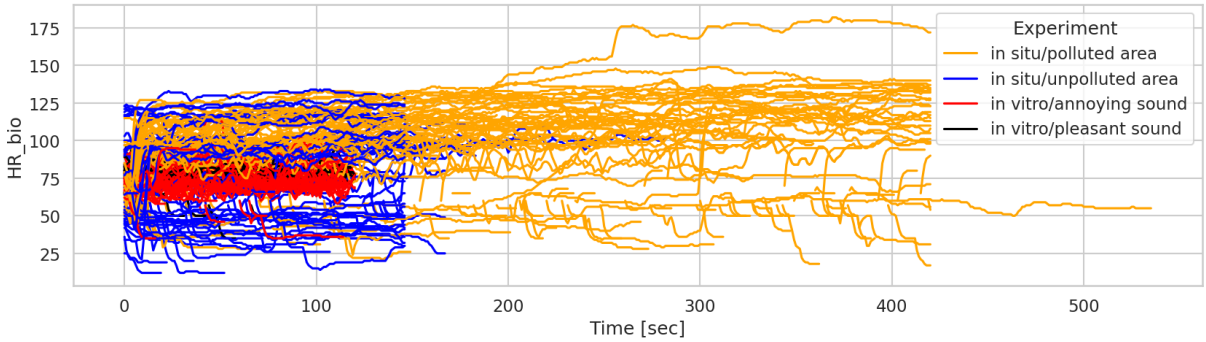


Figure 4.12: Refined heart rate data post-outlier removal: distribution and spread after data cleaning.

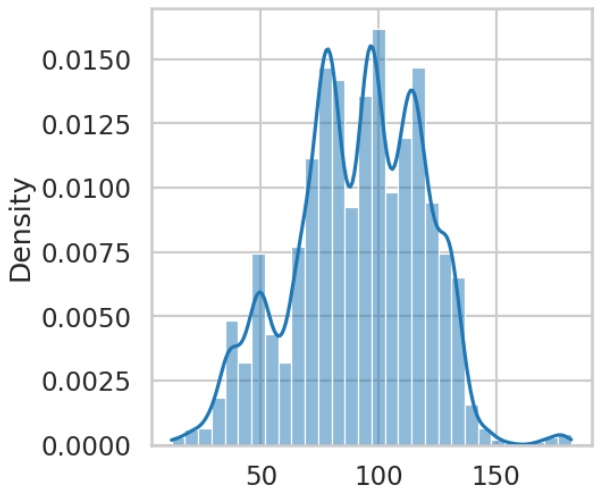


Figure 4.13: Distribution of heart rate data after outlier correction, showing reduced noise and more physiologically plausible values.

The extent of missing data was then assessed across all features and experimental conditions. The percentage of missing data was computed for each participant and experiment, and data streams with excessive loss (such as HR from the polluted environment, which exhibited significant gaps) were excluded from further analysis. See Figure 4.14 for an example visualisation of missingness across experiments.

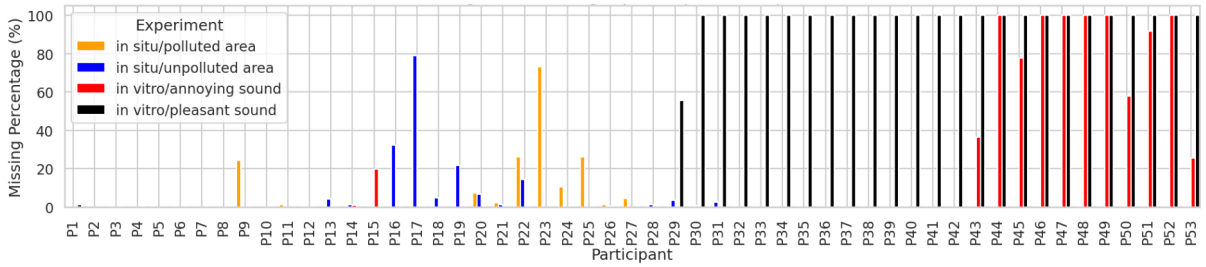


Figure 4.14: Heart rate data consistency and percentage of missing values across experiments and participants, informing feature selection for analysis.

After the data cleaning stage, any remaining missing values were addressed to preserve the continuity and integrity of the time-series data. For gaps occurring randomly within sequences, linear interpolation was applied to estimate plausible intermediate values based on neighbouring points. When missing values occurred at the beginning of a sequence, backward filling was used, propagating the first valid observation backwards to fill the gap. Conversely, for missing values at the end of a sequence, forward filling was employed, carrying the last valid value

forward. This combination of imputation methods ensured that the time-series data remained as complete and as representative as possible, supporting robust analytical outcomes in subsequent modelling and statistical analyses.

Through this rigorous, multi-step preprocessing pipeline, the EnviroWellBeing dataset was transformed into a synchronised, clean, and analysis-ready resource, optimised for exploring the interplay between environmental conditions and physiological responses. This section provides a foundation for the robust time-series and machine learning analyses detailed in subsequent chapters.

4.5 Preliminary Data Quality

Following the initial data preprocessing, a preliminary assessment of the EnviroWellBeing dataset was performed to evaluate the quality and variability of physiological and self-reported responses under different experimental scenarios.

Figure 4.15 presents the distribution of self-reported happiness levels (used as a proxy for stress) across the four experimental conditions. Lower happiness scores were observed in polluted and annoying sound environments, while higher scores were reported in unpolluted and pleasant sound conditions. These trends highlight the sensitivity of subjective wellbeing to air quality and auditory factors.

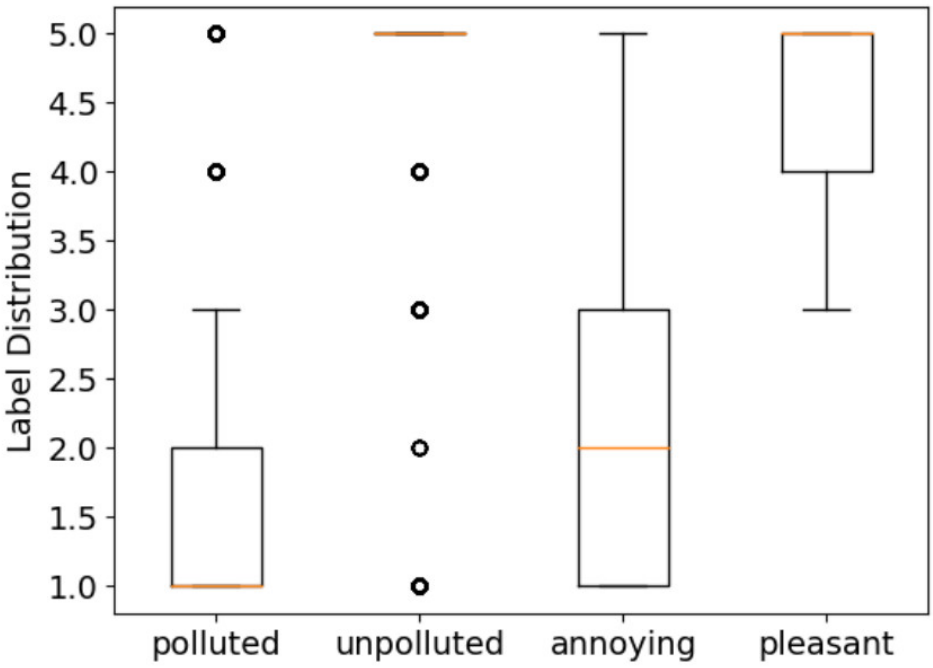


Figure 4.15: Self-reported happiness levels (proxy for stress) across polluted, unpolluted, annoying, and pleasant conditions.

Physiological data further demonstrate environmental influences on HR and breathing rate (BR). Figure 4.16 compares the distributions of HR and BR in polluted and unpolluted environments. Both HR and BR exhibited greater variability and generally higher values in polluted conditions, indicating increased physiological arousal.

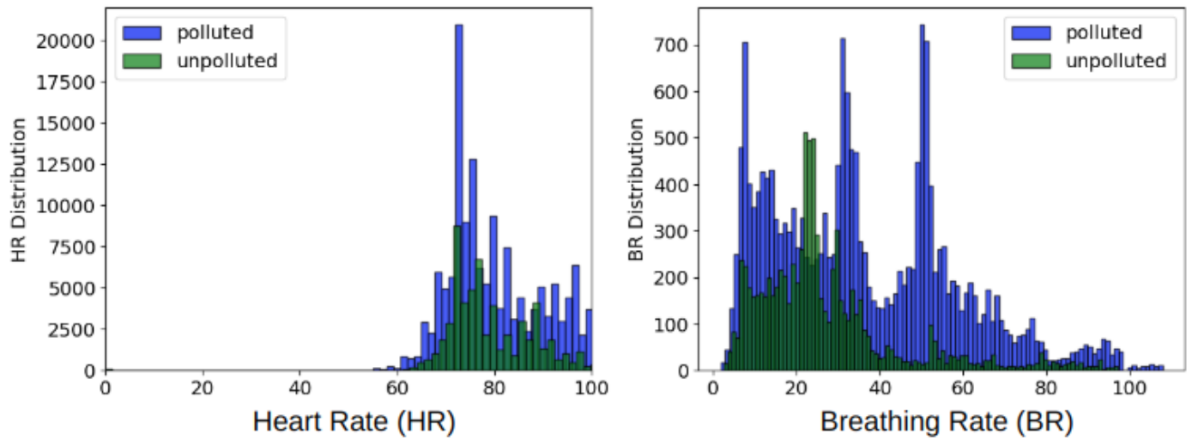


Figure 4.16: Distributions of heart rate (left) and breathing rate (right) in polluted and unpolluted environments.

For the sound exposure scenarios, Figure 4.17 shows distributions of HR and BR under annoy-

ing and pleasant auditory conditions. In this case, differences between conditions were comparatively modest, suggesting that acute auditory stimuli may have a less pronounced effect on these physiological measures than air quality.

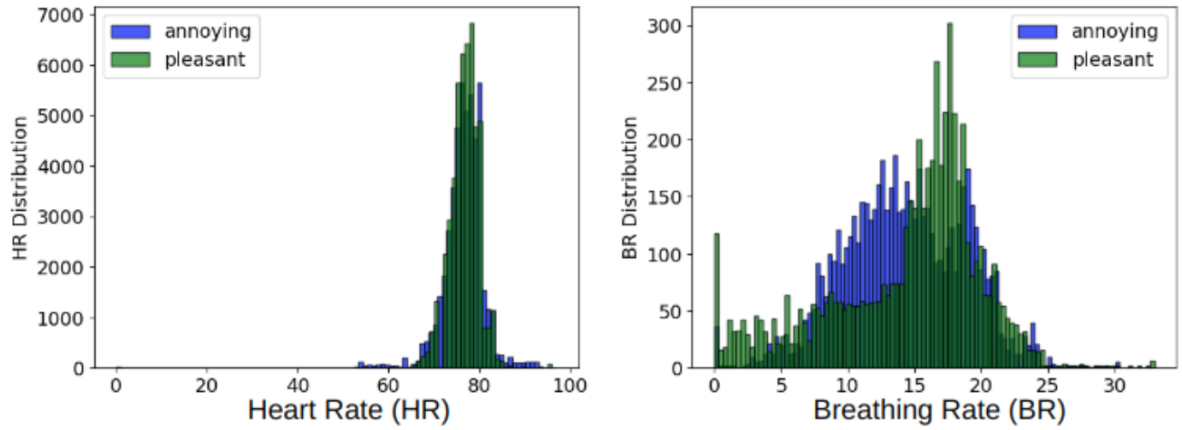


Figure 4.17: Distributions of heart rate (left) and breathing rate (right) in annoying and pleasant sound conditions.

This preliminary evaluation indicates that the EnviroWellBeing dataset effectively captures both subjective and physiological variation in response to environmental factors, supporting subsequent in-depth analyses.

To contextualise the quality and scope of the *EnviroWellBeing* dataset, it is helpful to compare it with prior efforts such as the *DigitalExposome* study by Johnson et al. [173]. While *DigitalExposome* explored the relationship between urban environments and wellbeing using real-time multi-sensor fusion with a smaller participant group (12 individuals after filtering), my dataset offers greater participant diversity (53 individuals), longer-term recordings, and richer time-series data. Moreover, unlike the exploratory nature of Johnson's work, my study integrates advanced preprocessing techniques such as PCA, feature selection, and statistical filtering to ensure high data fidelity. These enhancements not only contribute to improved model accuracy but also strengthen the reliability and applicability of the dataset for real-world use cases in mental health diagnostics and urban wellbeing analysis.

4.6 Questionnaire

Before starting the experiments, each participant completed a survey comprised of three questionnaires. The first was a demographic questionnaire please see appendix B.1 for further details , the second questionnaire aimed to measure the stress level of each participant using the Perceived Stress Scale (PSS-10) please see appendix B.3 for further details . The third was a short-form noise sensitivity questionnaire (NoiSeQ-SF) [174, 175] please see appendix B.2 for further details. The goal was to establish a baseline understanding of each participant's stress level and noise sensitivity prior to the experimental exposures.

4.6.1 Participants

Fifty-three participants completed both questionnaires. The cohort consisted of 32 males (60%) and 21 females (40%). The majority (n=42, 79%) were aged between 18-30 years. Most participants identified as White ethnicity (n=20, 38%) or Asian (n=15, 28%), were unmarried (n=39, 74%), and were students (n=37, 70%). Table 4.2 provides a detailed breakdown of the participant group characteristics.

4.6.2 Stress Level

The Perceived Stress Scale (PSS-10) [176], a widely used instrument, measured participants' stress levels. This scale consists of 10 questions regarding feelings and thoughts during the last month. Responses were given on a five-point scale from 'never' (0) to 'very often' (4). The total PSS score, ranging from 0 to 40, was calculated by summing the scores for all ten questions, with higher scores indicating higher levels of perceived stress. While not a clinical diagnostic tool, the PSS provides a good indication of stress levels. For descriptive purposes, scores were categorised: Low stress ($\leq$ 13), Moderate stress (14-26), and High stress ($<$ 26). As shown in Table 4.3, the majority of participants (n=46, 86.8%) fell into the moderate stress category based on their baseline scores.

Table 4.2: Participants group characteristics.

Category	Group	Number (n=53)	Percentage (%)
Gender	Male	32	60
	Female	21	40
Age	18-30	42	79
	30-40	6	11
	40-50	3	6
	50 and older	2	4
Ethnic origin	White	20	38
	Black	3	6
	Asian	15	28
	Other	15	28
Marital Status	Single	39	74
	Married	14	26
Education	High school	20	38
	College	5	9
	Bachelor's degree	15	28
	Master/PhD	13	25
Employment	Student	37	70
	Employed	12	22
	Unemployed	1	2
	Unable to work	3	6

Table 4.3: Stress level categories for all participants based on PSS-10 score.

Stress level	Number of participants	Percentage (%)
Low ($\leq$ 13)	6	11.4
Moderate (14-26)	46	86.8
High ($>$ 26)	1	1.8

4.6.3 Noise Sensitivity

Participants' sensitivity to noise was measured using the 15-item Noise Sensitivity Questionnaire Short Form (NoiSeQ-SF), derived from the original NoiSeQ [174, 175]. This short form provides a reliable estimate of global noise sensitivity for normally hearing persons. Responses were recorded on a 4-point Likert scale from "strongly disagree" (1) to "strongly agree" (4), with higher total scores (sum of item scores, range 15-60) indicating a higher degree of noise sensitivity [175]. For descriptive analysis, scores were categorised: Low Noise Sensitivity ($\leq$ 20), Moderate Noise Sensitivity (21-40), and High Noise Sensitivity ($\geq$ 40). The results, sum-

marised in Table 4.4, indicate that most participants exhibited moderate (n=43, 81.1%) or high (n=10, 18.9%) noise sensitivity at baseline.

Table 4.4: Noise sensitivity level categories for all participants based on NoiSeQ-SF score.

NS level	Number of participants	Percentage (%)
Low ($\leq$ 20)	0	0
Moderate (21-40)	43	81.1
High ($\geq$ 40)	10	18.9

4.7 Summary

This chapter introduced the EnviroWellBeing Dataset, developed to address gaps in existing environmental health research. Key aspects included the rationale for dataset creation, an overview of the wearable and environmental sensors used, experimental protocols for data collection, and the main preprocessing steps applied to the data. The chapter also provided a preliminary evaluation of data quality and summarised participants' baseline stress and noise sensitivity levels. These foundations support the advanced analyses presented in the following chapter 5, where machine learning methods are applied to explore environmental impacts on wellbeing.

Chapter 5

Environmental Impacts on Wellbeing

5.1 Introduction

This chapter delves into the core analyses of the EnviroWellBeing dataset (Chapter 4), addressing key research questions regarding the environmental impact on stress and the value of different analytical approaches. Specifically, it investigates the impact of environmental conditions on stress (RQ2), the nature of this impact by examining predictive features (RQ3), and the contribution of time-series data compared to snapshot information (RQ4). Table 5.1 presents an overview of the physiological and environmental features utilised in the different classification tasks performed in this chapter.

Table 5.1: Feature Selection for Environmental and Stress Level Classification Tasks

Classification type	**Exp. Type**	BR	Posture	Activity	Acceleration	BR Amplitude	ECG Amplitude	ECG Noise	RR	EDA	HR wrist	Temp wrist	UV	EnvNoise	AirPressure	Temp dl	Light	Sound	VOC	CO	CO2	NO	NO2	SO2
Environmental classification	In Situ	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗
	In Vitro	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗
Stress level classification	In Situ	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
	In Vitro	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✗	✗	✗	✗	✗	✗	✗	✗	✗

The analyses employ LSTM and 1D-CNN based deep learning techniques for time-series classification to distinguish environmental conditions using only physiological data. Furthermore, the predictive capability of combining environmental and physiological data to estimate self-

reported stress levels is explored using RF, LSTM, and 1D-CNN models. Successfully classifying environments and predicting stress from sensor data holds significant real-world implications, potentially enabling personalised environmental health feedback systems, objective stress monitoring tools, and providing evidence to inform urban design strategies aimed at enhancing public wellbeing. The subsequent sections detail the data preparation, methodology, and results for these analyses.

5.2 Data Preparation

This section details the data preparation pipeline applied to the preprocessed EnviroWellBeing dataset prior to the classification analyses. The aim was to structure the data for both time-series (sequential) and snapshot (non-sequential) modelling, ensuring comparability between analytic approaches and supporting robust evaluation of each model's performance. The specific feature sets for each task are summarised in Table 5.1.

5.2.1 Data Normalisation

To ensure all features contributed comparably during model training, and to enhance convergence and stability, data normalisation was applied as a standard preprocessing step. The **Normalizer** function from **sklearn.preprocessing** was used, scaling each sample to unit L2 norm [165]. This approach is particularly suitable for multimodal datasets with varying signal magnitudes. Table 5.2 summarises several common normalisation techniques for context.

Table 5.2: Common Data Normalisation Techniques

Normalisation Technique	Equation	Typical Use Case
Linear Scaling (Min-Max)	$x' = (x - x_{min})/(x_{max} - x_{min})$	Uniformly distributed features
Clipping	if $x >$ max, then $x' =$ max. if $x <$ min, then $x' =$ min	Handling extreme outliers
Log Scaling	$x' = \log(x)$	Features following power law
Z-score Standardisation	$x' = (x - \mu)/\sigma$	Features without extreme outliers

5.2.2 Data Segmentation

For sequence-based models (LSTM, 1D-CNN), the continuous time-series data (synchronised at 1Hz, see Chapter 4) was segmented using a sliding window approach. Each segment comprised 30 seconds (30 time steps), with a step size of 10 seconds, resulting in a 20-second overlap between consecutive segments. This ensured sufficient temporal context for each sample and increased the effective number of training examples (see Figure 5.1).

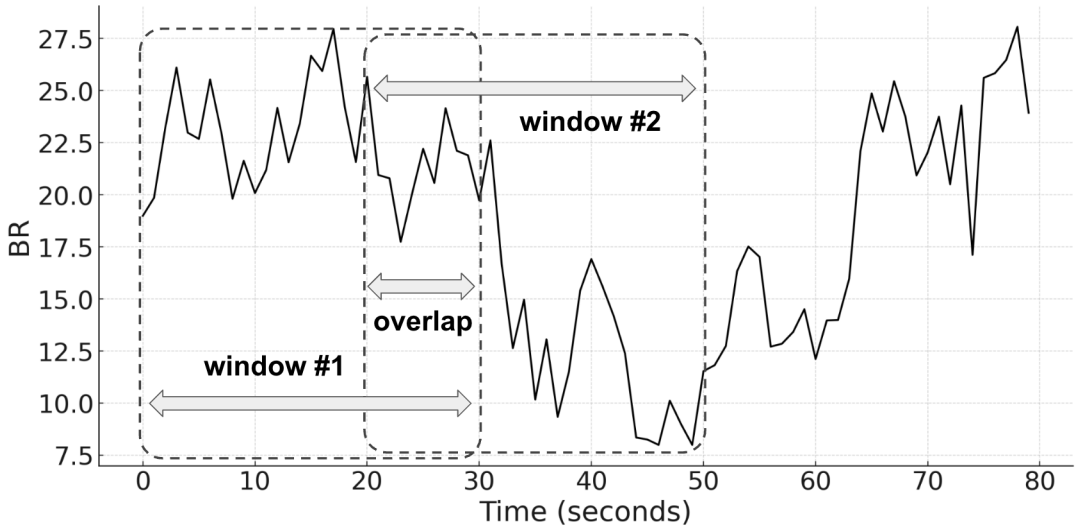


Figure 5.1: Time-series Data Segmentation using Sliding Window Technique

For snapshot-based (Random Forest) analysis, only the final time point of each 30-second segment was used as input, explicitly removing temporal dependencies and allowing direct comparison with time-series approaches.

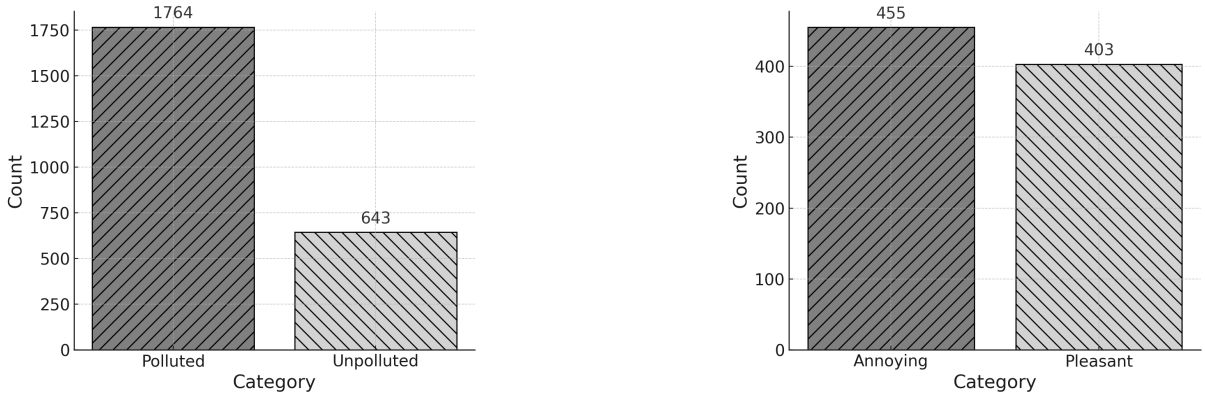


Figure 5.2: Sample Count per Class for Environment Classification: (Left) In Situ Experiments, (Right) In Vitro Experiments.

For the environment classification task, each segment was labelled according to the experimental condition, with '0' representing polluted or annoying environments and '1' representing unpolluted or pleasant environments (see Figure 5.2 for class distributions). In the stress level classification task, segments were labelled with the self-reported stress level (ranging from 1 to 5) recorded at the end of each segment (see Figure 5.3).

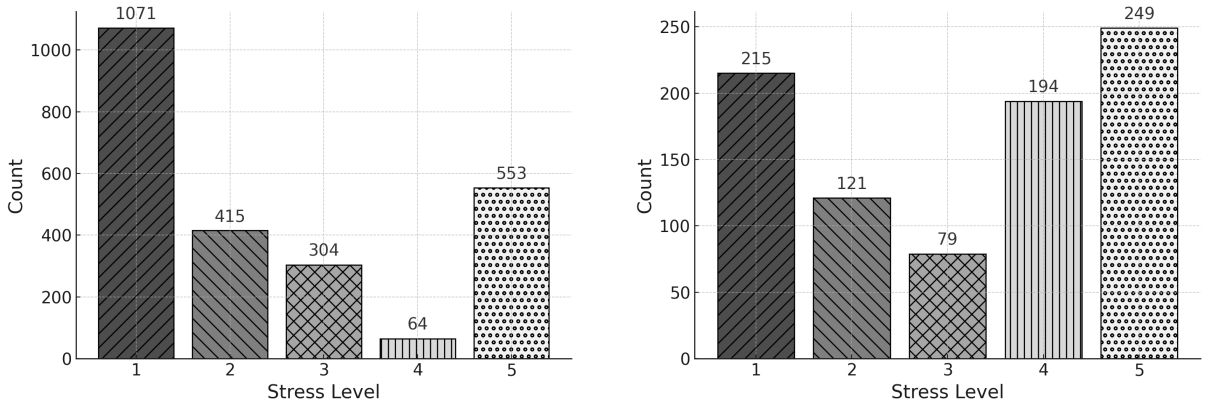


Figure 5.3: Sample Count per Class for Self-Reported Stress Level Classification: (Left) In Situ Experiments, (Right) In Vitro Experiments.

5.2.3 Train-Validation-Test Split

To robustly evaluate model generalisability and avoid data leakage, the prepared data for each task was stratified and split into training, validation, and test sets (60% / 20% / 20%), maintaining class distributions using **StratifiedKFold** from **scikit-learn** [165]. The conceptual overview is depicted in Figure 5.4.

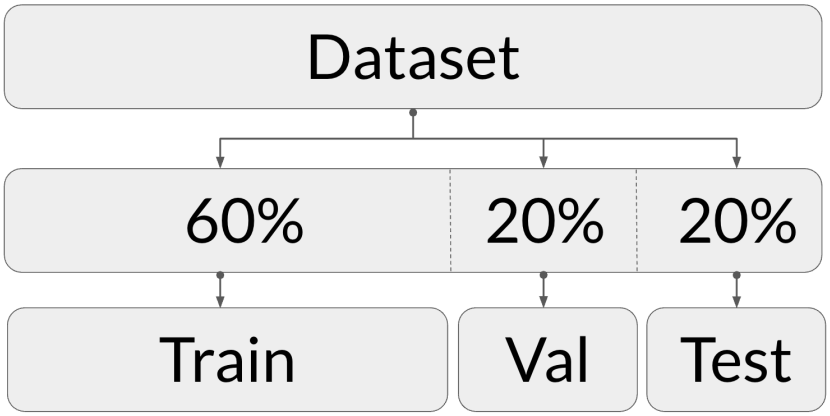


Figure 5.4: Train-Validation-Test Data Splitting Concept

Figures 5.5 and 5.6 show the class distributions across splits for both classification tasks.

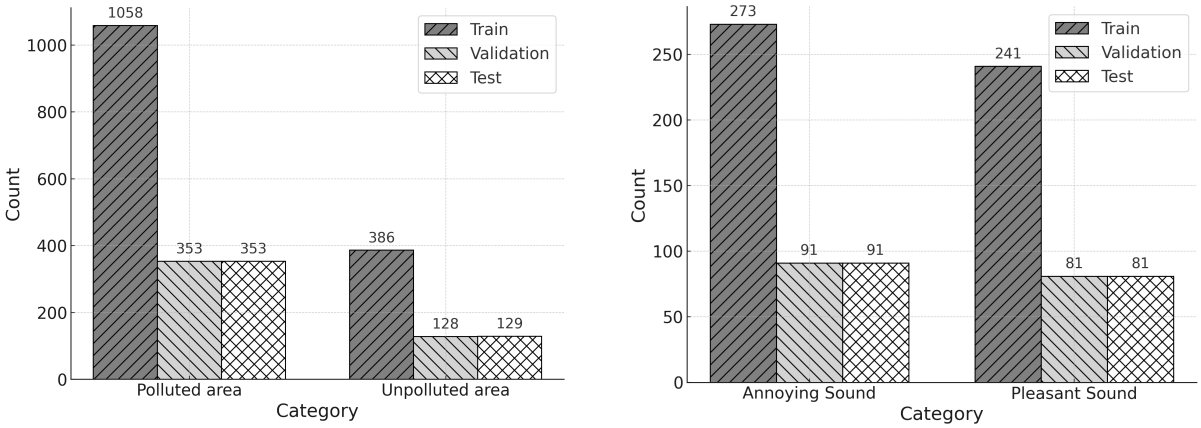


Figure 5.5: Sample Count per Class Across Train-Validation-Test for Environment Classification: (Left) In Situ Experiments, (Right) In Vitro Experiments

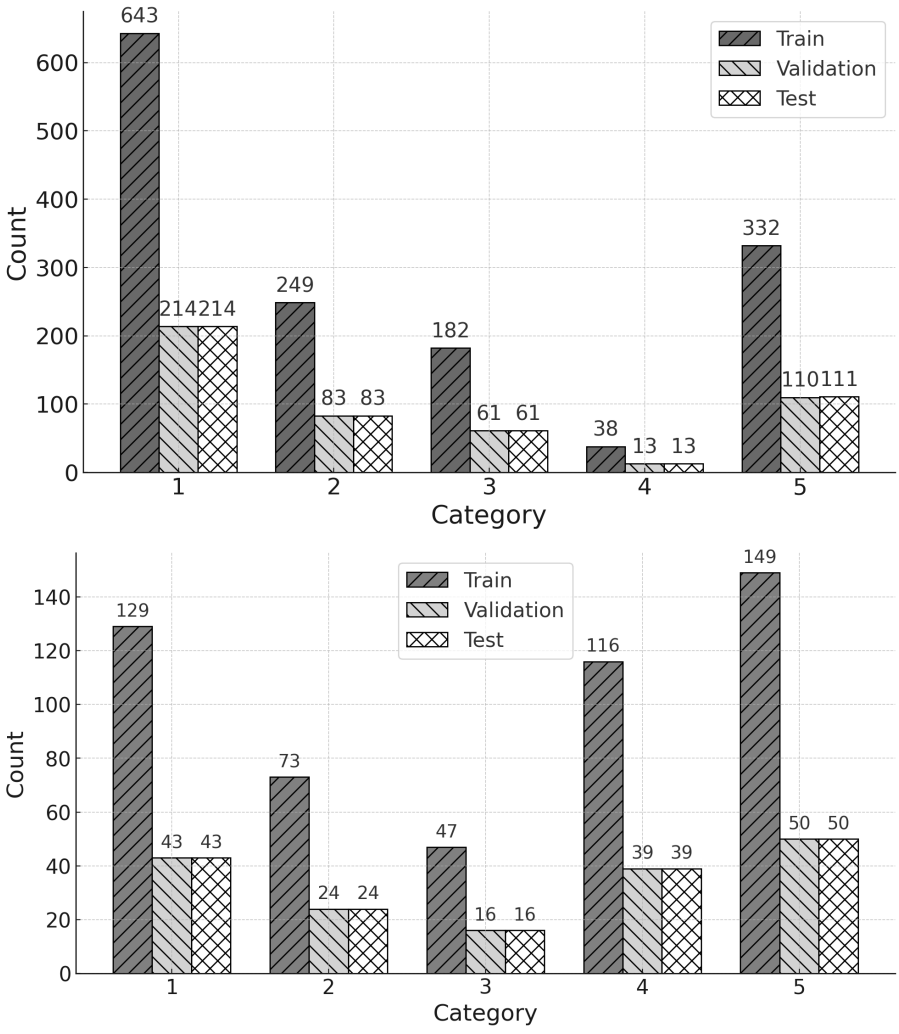


Figure 5.6: Sample Count per Class Across Train-Validation-Test for Stress Level Classification: (Top) In Situ Experiments, (Bottom) In Vitro Experiments

Table 5.3 provides a consolidated summary of the final dataset characteristics used for the experiments in this chapter, detailing sample counts, features, shapes, labels, and split sizes for each classification task.

Table 5.3: Overview of Data Segmentation and Classification Experiments

Classification Type	Experiment Type	Number of samples	Segment size	Number of selected features	Input shape	Labels	Train shape	Validation shape	Test shape
Environment Classification	In Situ	2407	30	11	(2407,30,11)	0-Polluted 1-Unpolluted	(1444,30,11)	(481,30,11)	(482,30,11)
	In Vitro	858	30	11	(858,30,11)	0-Annoying 1-Pleasant	(514,30,11)	(172,30,11)	(172,30,11)
Self-reported Stress Level Classification	In Situ	2407	30	23	(2407,30,23)	1-High Stress 2-Mild Stress 3-Neutral 4-Mild Relaxation 5-Relaxed	(1444,30,23)	(481,30,23)	(482,30,23)
	In Vitro	858	30	14	(858,30,14)	1-High Stress 2-Mild Stress 3-Neutral 4-Mild Relaxation 5-Relaxed	(514,30,14)	(172,30,14)	(172,30,14)

This preparation facilitated a fair comparison of sequential (LSTM, 1D-CNN) and non-sequential (RF) models in subsequent analyses, isolating the added value of temporal context in predicting both environmental conditions and human stress responses.

5.3 Methodology

This section outlines the methodology used to analyse the prepared EnviroWellBeing dataset for environmental classification and stress level prediction. The approach leverages both time-series models (LSTM, 1D-CNN) and a snapshot-based model (Random Forest), enabling a direct comparison of sequential versus non-sequential methods for predicting environmental context and stress from multimodal sensor data. The subsequent subsections describe the overall system workflow, model architectures, and the evaluation metrics employed.

5.3.1 Modelling Flowchart

The overall workflow, summarised in Figure 5.7, follows a standard machine learning pipeline. Data is preprocessed and segmented (Section 5.2), then fed into the appropriate models for each analytic task. Deep learning architectures (LSTM and 1D-CNN) are designed to cap-

ture temporal dependencies within the 30-second physiological and environmental sequences, while Random Forest is applied to instantaneous (last-second) snapshots of the same features. This side-by-side design enables direct evaluation of the added value of temporal modelling, as models can be compared under identical data and feature conditions. Feature engineering is handled within the deep learning models, and hyperparameter optimization is conducted based on validation set performance.

The flowchart presents the sequence of steps from data input through to model training and final evaluation, and is streamlined to include only the methods and metrics implemented in this study, reflecting revisions made per examiner feedback.

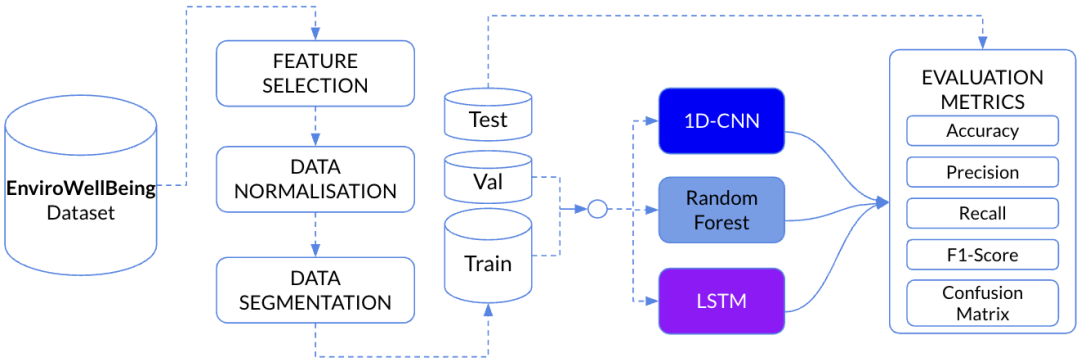


Figure 5.7: Modelling flowchart for environmental and stress level classification tasks.

5.3.2 Model Architectures

The LSTM network consists of a single layer with 64 units, followed by a dropout layer (dropout rate 0.3) to mitigate overfitting, and a final dense output layer for classification. The 1D-CNN model includes two convolutional layers (each with 64 filters, ReLU activation), followed by a dense layer and softmax output. Both models are trained using the Adam optimizer, with loss functions adapted to the task: binary cross-entropy for environment classification, and categorical cross-entropy for multi-class stress level prediction. The Random Forest classifier serves as a snapshot baseline, using the last time step of each segment as input. Model hyperparameters (such as number of units/filters) were selected empirically based on validation performance.

5.3.3 Evaluation Metrics

To assess the performance of the classification models, standard evaluation metrics derived from the confusion matrix were used. The confusion matrix summarises the counts of True Positives (TP - correctly predicted positive instances), True Negatives (TN - correctly predicted negative instances), False Positives (FP - incorrectly predicted positive instances, Type I error), and False Negatives (FN - incorrectly predicted negative instances, Type II error).

Based on these counts, the following metrics were calculated:

Accuracy: The proportion of total predictions that were correct.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{5.1}$$

Precision: The proportion of positive predictions that were actually correct. It measures the exactness of the classifier.

$$Precision = \frac{TP}{TP + FP} \tag{5.2}$$

Recall (Sensitivity): The proportion of actual positive instances that were correctly identified. It measures the completeness of the classifier.

$$Recall = \frac{TP}{TP + FN} \tag{5.3}$$

F1-Score: The harmonic mean of Precision and Recall, providing a single score that balances both metrics. It is useful especially when class distribution is uneven.

$$F1\text{-}Score = 2 \times \frac{Precision \times Recall}{Precision + Recall} = \frac{2TP}{2TP + FP + FN} \tag{5.4}$$

For multi-class problems like stress level prediction (1-5), these metrics are typically calculated for each class individually (treating it as the 'positive' class against all others) and then averaged (e.g., macro-average or weighted-average) to obtain overall performance scores.

5.4 Results

This section presents a comparative analysis of the performances of LSTM and 1D-CNN networks applied to the tasks of Environmental Classification and Self-Reported Stress Level Classification using the prepared EnviroWellBeing dataset. The evaluation is performed separately for the ”In Situ” and ”In Vitro” experiments using metrics including accuracy, precision, recall, and F1 score, derived from the performance of the test set.

5.4.1 Environment Classification

The Environment Classification task aimed to classify the environment type (Polluted vs. Unpolluted) based solely on physiological data features (listed in Table 5.1). LSTM and 1D-CNN models were evaluated.

5.4.1.1 Environment Classification: In Situ

The ’In Situ’ environmental classification utilised physiological data to discern between ’Polluted’ and ’Unpolluted’ areas. Figures 5.8 and 5.9 display the loss and accuracy visualisations over 100 epochs for both models.

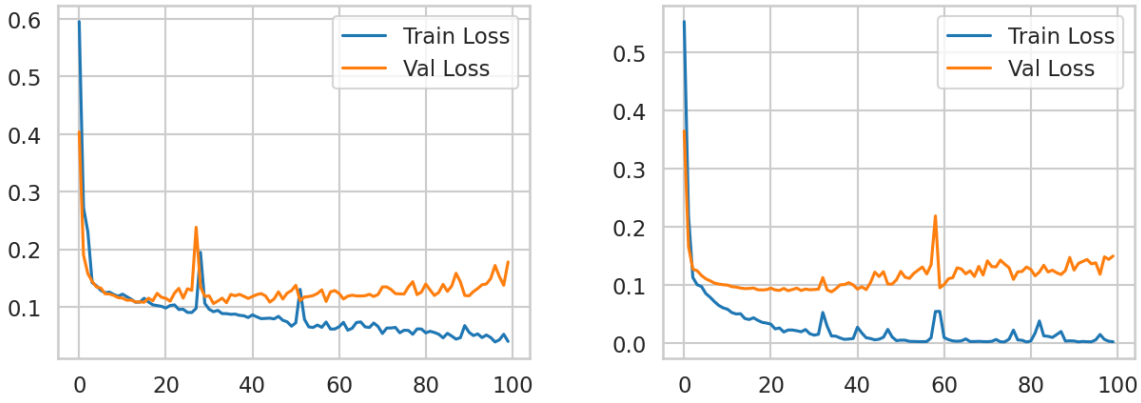


Figure 5.8: Training and validation loss curves for (left) LSTM and (right) 1D-CNN models during In Situ environment classification.

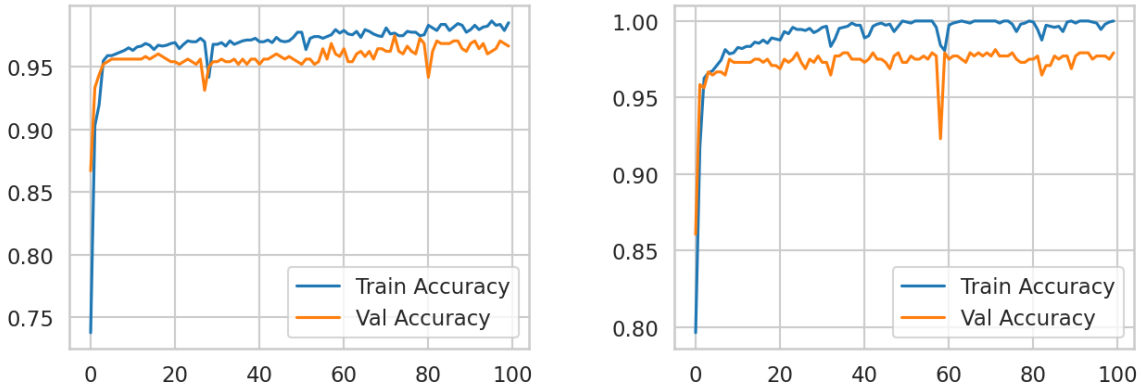


Figure 5.9: Training and validation accuracy curves for (left) LSTM and (right) 1D-CNN models during In Situ environment classification.

For the LSTM model, the training curves showed stable learning. The 1D-CNN model exhibited smoother convergence and slightly higher validation accuracy. The confusion matrices for the validation and test datasets offer a depiction of predictive accuracy (Figures 5.10, 5.11, 5.12, and 5.13).

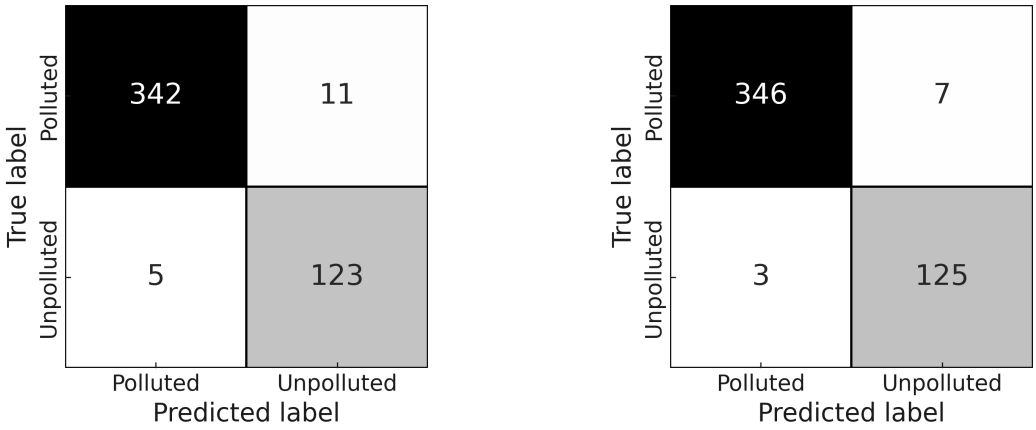


Figure 5.10: Validation set confusion matrices for (left) LSTM and (right) 1D-CNN models in In Situ environment classification.

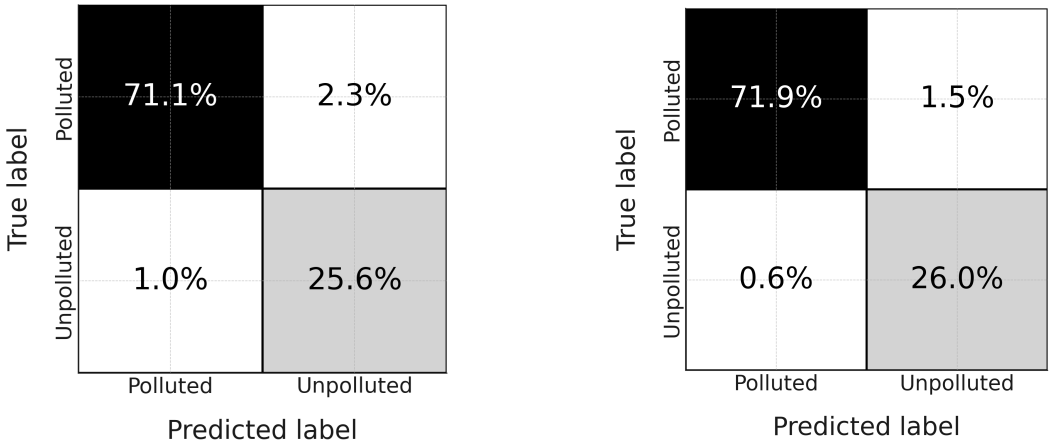


Figure 5.11: Validation set normalised confusion matrices for (left) LSTM and (right) 1D-CNN models in In Situ environment classification.

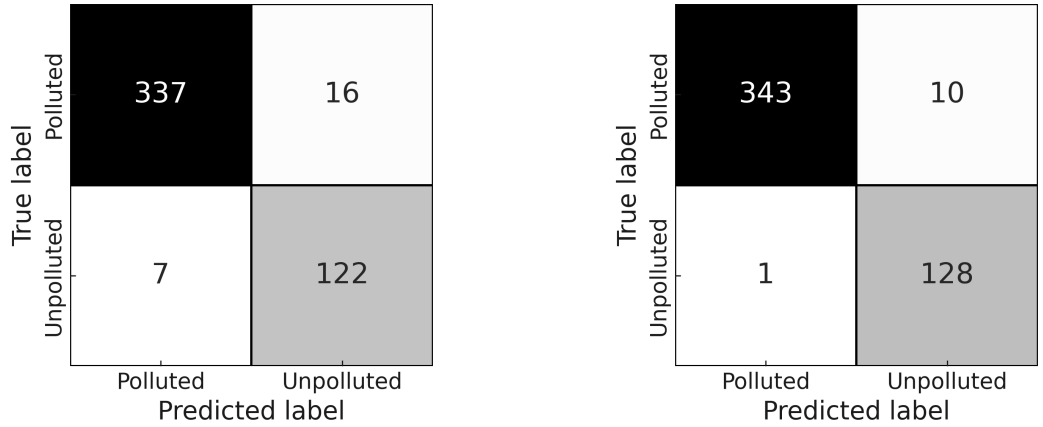


Figure 5.12: Test set confusion matrices for (left) LSTM and (right) 1D-CNN models in In Situ environment classification.

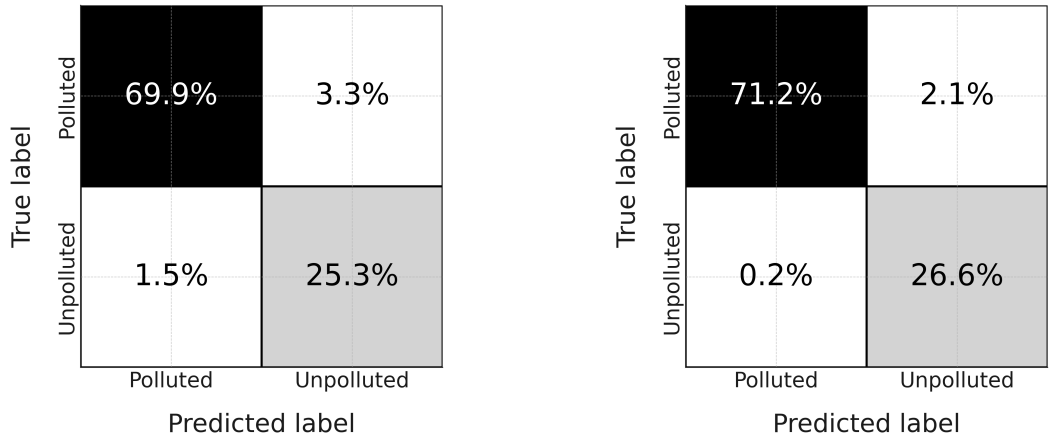


Figure 5.13: Test set normalised confusion matrices for (left) LSTM and (right) 1D-CNN models in In Situ environment classification.

The 1D-CNN model consistently showed superior performance with fewer misclassifications on both validation and test sets. Table 5.4 quantitatively summarises the performance metrics for both models.

Table 5.4: Environment Classification (in Situ): LSTM vs. 1D-CNN (All Values in %)

Model	Polluted			Unpolluted			Accuracy
	Precision	Recall	F1-Score	Precision	Recall	F1-Score	
LSTM (Validation)	98.56↓	96.88↓	97.71↓	91.79↓	96.09↓	93.89↓	96.67↓
LSTM (Test)	97.97↓	95.47↓	96.70↓	88.41↓	94.57↓	91.39↓	95.23↓
1D-CNN (Validation)	**99.14↑**	**98.02↑**	**98.58↑**	**94.70↑**	**97.66↑**	**96.15↑**	**97.92↑**
1D-CNN (Test)	**99.71↑**	**97.17↑**	**98.42↑**	**92.75↑**	**99.22↑**	**95.88↑**	**97.72↑**

The 1D-CNN outperformed the LSTM across all metrics, achieving a test accuracy of 97.72% compared to 95.23% for LSTM.

5.4.1.2 Environment Classification: In Vitro

The 'In Vitro' environmental classification involved discerning between 'Annoying' and 'Pleasant' sound conditions based on physiological data responses. The LSTM and 1D-CNN models underwent a rigorous evaluation through training, validation, and testing phases. Figures 5.14 and 5.15 showcase the model loss and accuracy visualisations.

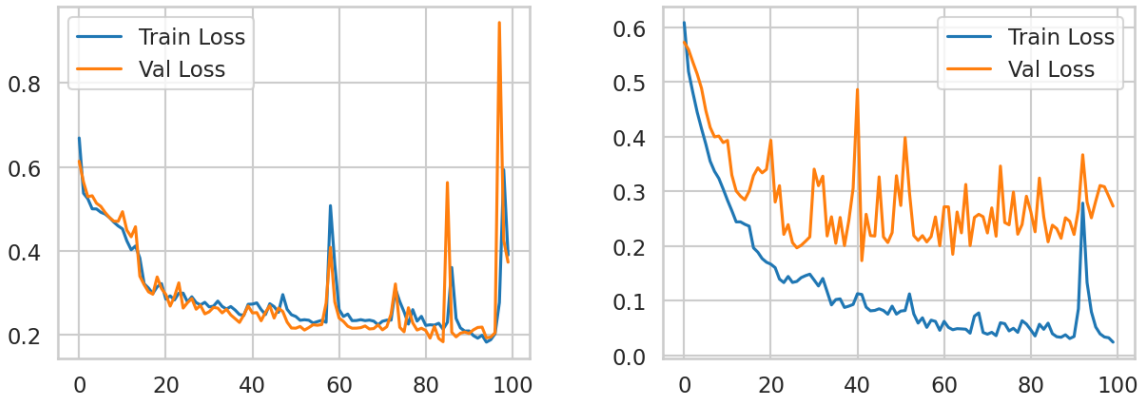


Figure 5.14: Training and validation loss curves for (left) LSTM and (right) 1D-CNN models during In Vitro environment classification.

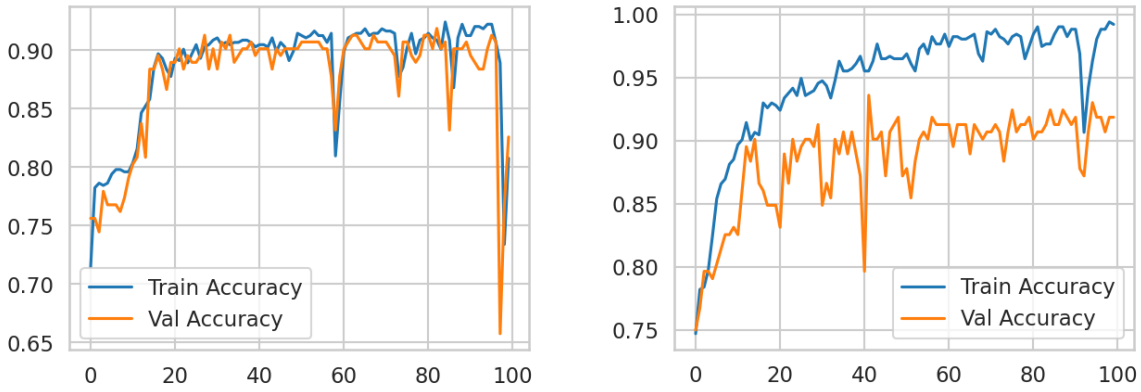


Figure 5.15: Training and validation accuracy curves for (left) LSTM and (right) 1D-CNN models in In Vitro environment classification.

For the LSTM model, Figure 5.14(left) illustrates fluctuating training and validation loss. Despite this, the LSTM model managed to achieve reasonable accuracy (Figure 5.15(left)), although with variance in validation accuracy. Conversely, the 1D-CNN model depicted in Figure 5.14(right) indicates a smoother convergence in loss, and Figure 5.15(right) presents consistent high performance with good generalisation. The confusion matrices for the validation and test data provide further depth (Figures 5.16, 5.17, 5.18, and 5.19).

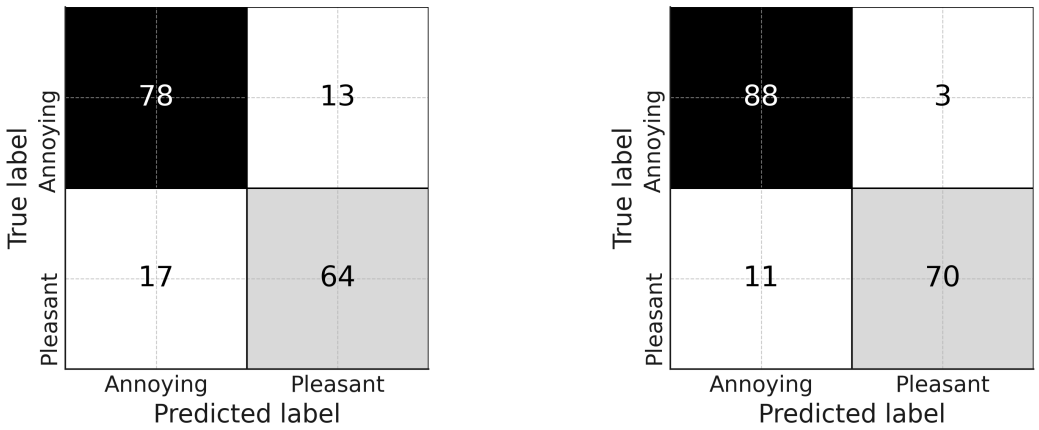


Figure 5.16: Validation set confusion matrices for (left) LSTM and (right) 1D-CNN models in In Vitro environment classification.

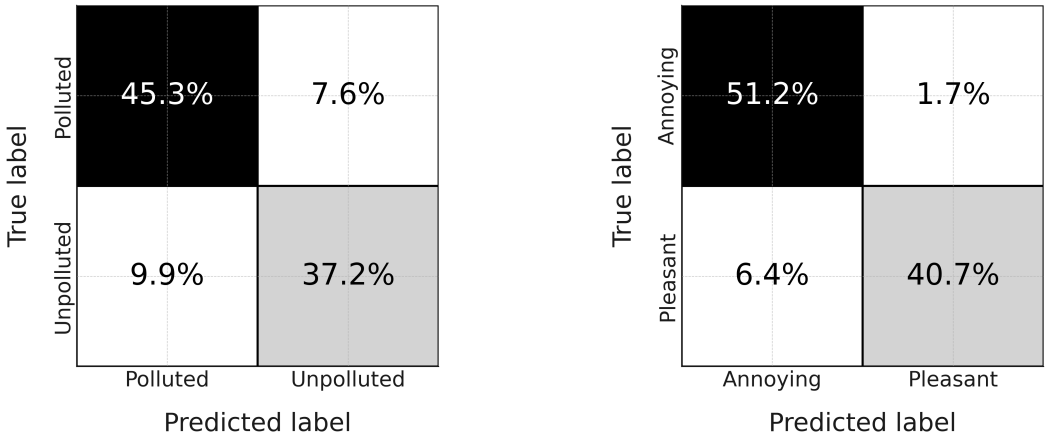


Figure 5.17: Validation set normalised confusion matrices for (left) LSTM and (right) 1D-CNN models in In Vitro environment classification.

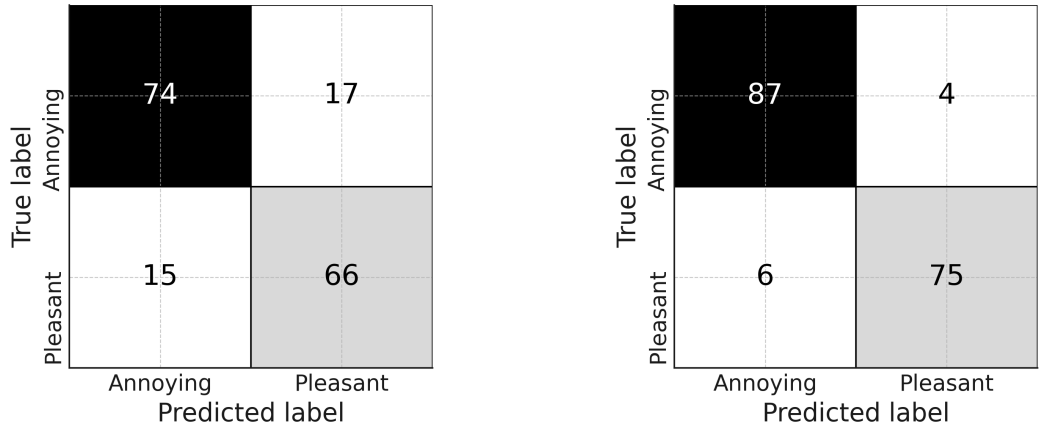


Figure 5.18: Test set confusion matrices for (left) LSTM and (right) 1D-CNN models in In Vitro environment classification.

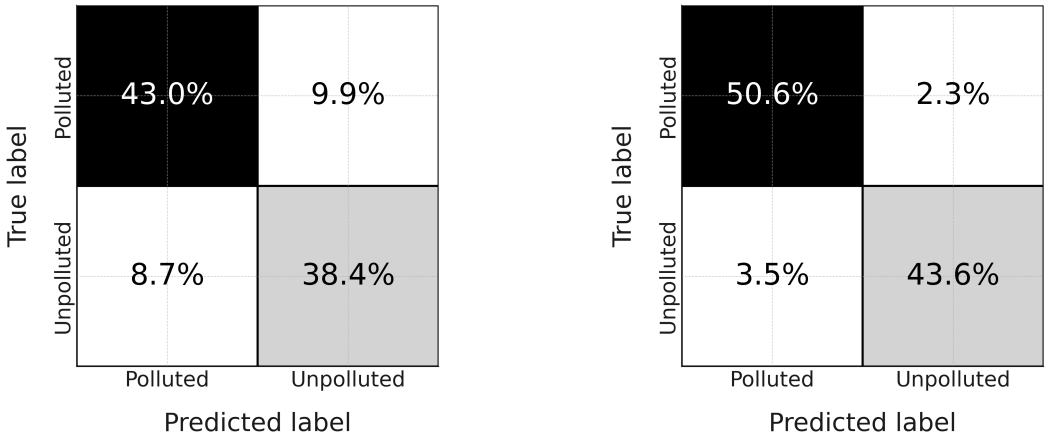


Figure 5.19: Test set confusion matrices for (left) LSTM and (right) 1D-CNN models in In Vitro environment classification.

The 1D-CNN model showed a marked improvement over LSTM in the confusion matrices. Table 5.5 details the quantitative comparison, formatted using the consistent ‘booktabs‘ style based on the original table structure.

Table 5.5: Environment Classification (in Vitro): LSTM vs. 1D-CNN (All Values in %)

Model	Annoying			Pleasant			Accuracy
	Precision	Recall	F1-Score	Precision	Recall	F1-Score	
LSTM (Validation)	82.11↓	85.71↓	83.87↓	83.12↓	79.01↓	81.01↓	82.56↓
LSTM (Test)	83.15↓	81.32↓	82.22↓	79.52↓	81.48↓	80.49↓	81.40↓
1D-CNN (Validation)	**88.89↑**	**96.70↑**	**92.63↑**	**95.89↑**	**86.42↑**	**90.91↑**	**91.86↑**
1D-CNN (Test)	**93.55↑**	**95.60↑**	**94.57↑**	**94.94↑**	**92.59↑**	**93.75↑**	**94.19↑**

The 1D-CNN model achieved higher test accuracy (94.18%) compared to LSTM (81.39%), with superior average precision, recall, and F1-scores, confirming its robustness for this task as well.

5.4.2 Self-Reported Stress Level Classification

5.4.2.1 Stress Level Classification: In Situ

In this section, the performance of Random Forest, LSTM, and 1D-CNN models is compared for the classification of self-reported stress levels (1–5) using the full set of physiological and environmental features (see Table 5.1) in real-world, in situ settings. LSTM and 1D-CNN utilise full 30-second time-series, while RF operates on the last second of each segment to represent snapshot-based analysis.

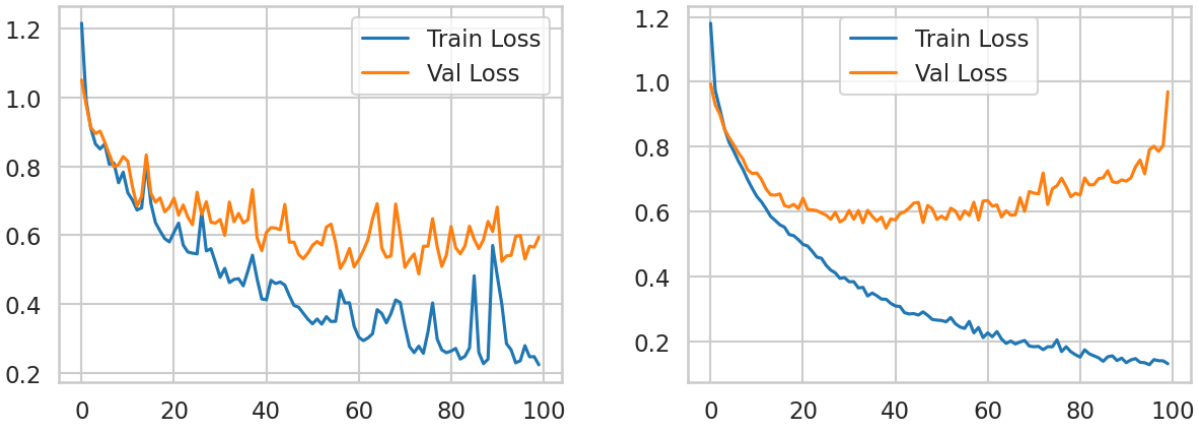


Figure 5.20: Training and validation loss curves for LSTM (left) and 1D-CNN (right) models in self-reported stress level classification (In Situ).

The training process and convergence of the deep learning models are visualized in the loss curves in Figure 5.20. The LSTM and 1D-CNN models both show decreasing validation and training loss over epochs, with 1D-CNN typically exhibiting smoother convergence.

Validation and training accuracy curves for LSTM and 1D-CNN models provide further insight into model generalization. Both models achieve increasing accuracy over epochs, with 1D-CNN generally maintaining more stable accuracy (see Figure 5.21).

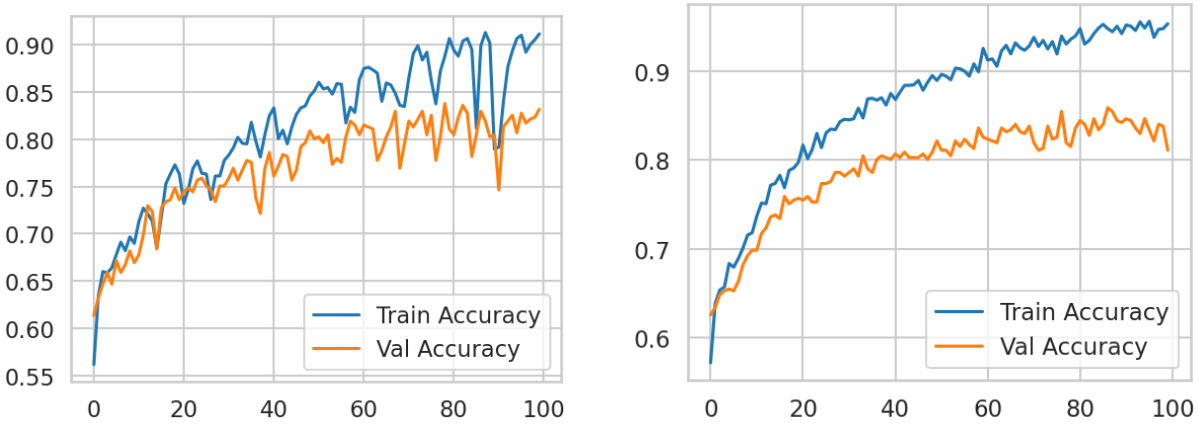


Figure 5.21: Training and validation accuracy curves for LSTM (left) and 1D-CNN (right) models in self-reported stress level classification (In Situ).

Model performance on the validation set is illustrated by the confusion matrices and their normalisations in Figure 5.22 and 5.23: RF (left), LSTM (center), and 1D-CNN (right). These matrices reveal each model's ability to correctly predict stress levels, highlighting that 1D-CNN

and LSTM tend to achieve higher accuracy and fewer confusions between stress categories compared to RF.

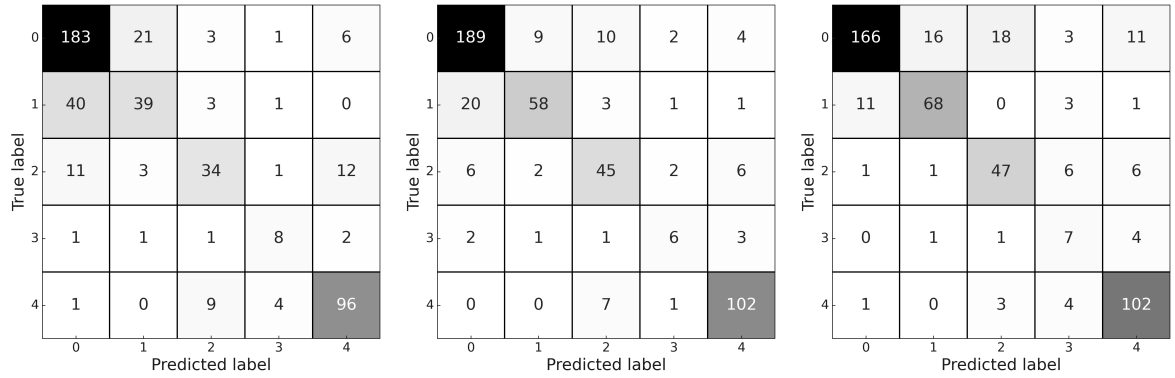


Figure 5.22: Validation confusion matrices for self-reported stress level classification (RF: left, LSTM: centre, 1D-CNN: right) in In Situ experiments.

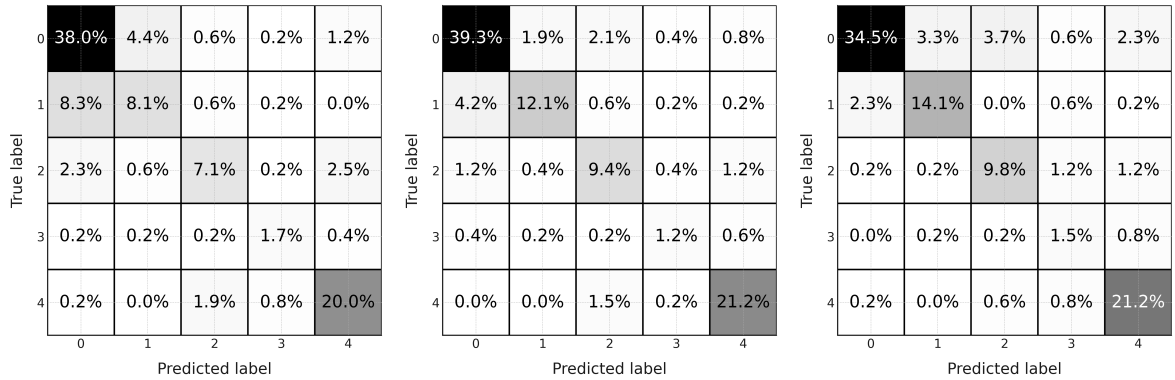


Figure 5.23: Normalised validation confusion matrices for self-reported stress level classification (RF: left, LSTM: centre, 1D-CNN: right) in In Situ experiments.

The Figure 5.24 and 5.25 shows confusion matrices and normalised confusion matrices for the test set results of each model: RF (left), LSTM (centre), and 1D-CNN (right). The 1D-CNN maintains the highest overall performance, particularly in correctly identifying both low and high stress levels, whereas RF shows more confusion between mid-range stress levels.

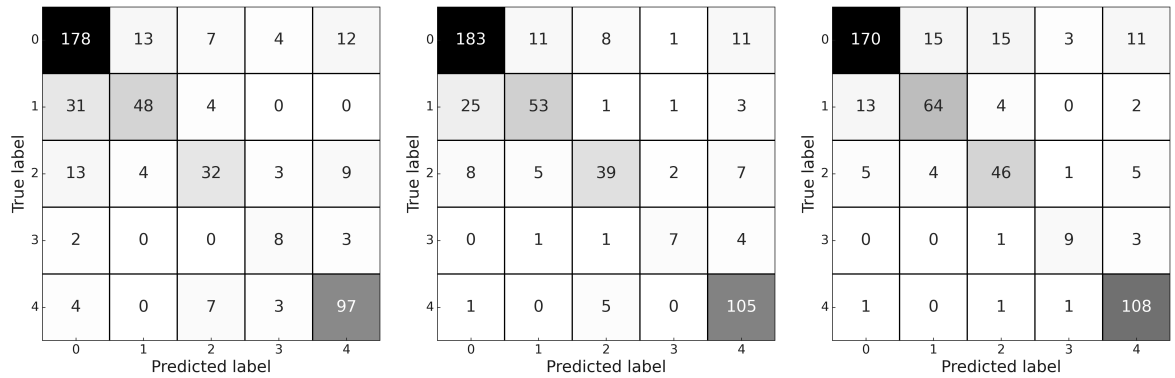


Figure 5.24: Test confusion matrices for self-reported stress level classification (RF: left, LSTM: centre, 1D-CNN: right) in In Situ experiments.

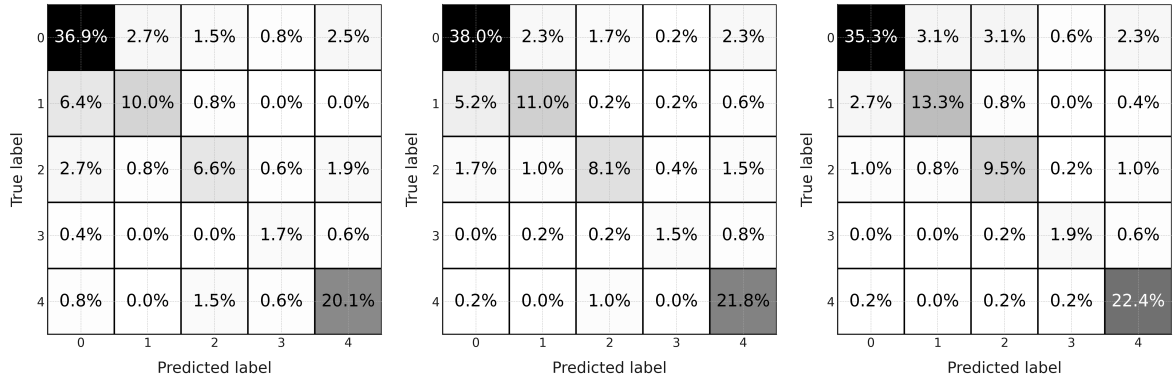


Figure 5.25: Normalised test confusion matrices for self-reported stress level classification (RF: left, LSTM: centre, 1D-CNN: right) in In Situ experiments.

The Random Forest model provides insight into which features are most informative for predicting self-reported stress levels (see Figure 5.26). In this study, environmental factors such as CO_2 and NO concentration, along with physiological signals like wrist temperature, EDA, and RR interval, were consistently identified as key contributors.

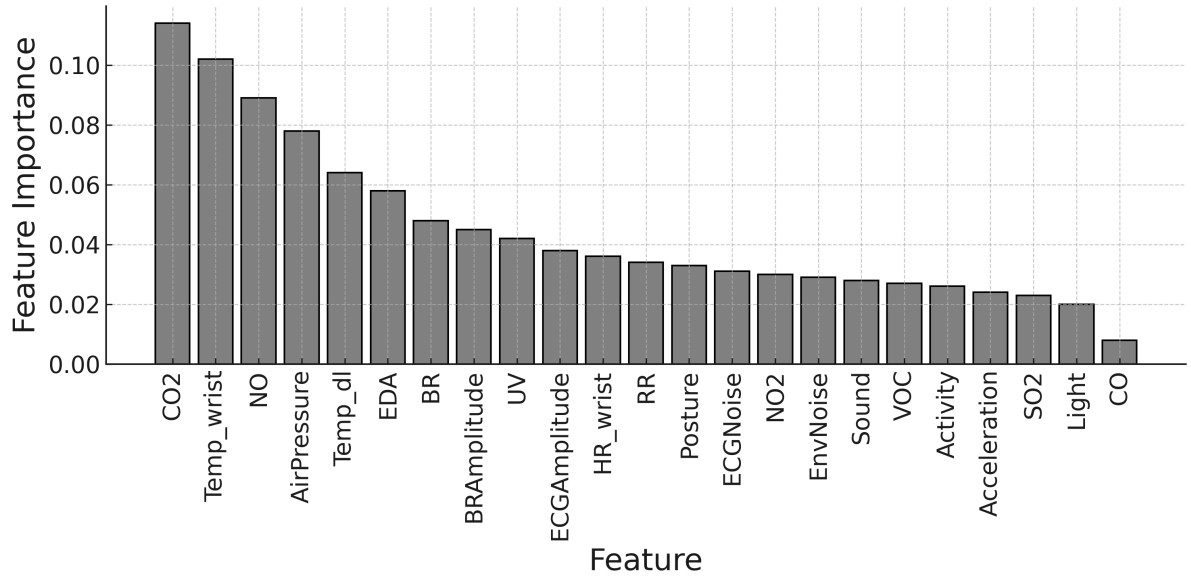


Figure 5.26: Feature importance from Random Forest classifier for self-reported stress level classification in In Situ experiments.

A summary of model performance for each stress level, including validation and test metrics (accuracy, precision, recall, F1), is presented in Table 5.6. The 1D-CNN model achieved the highest average test accuracy (82.37%), outperforming both LSTM and Random Forest, underscoring the benefit of leveraging temporal dynamics in physiological and environmental data.

Table 5.6: Stress Level Classification (In Situ): Random Forest vs. LSTM vs. 1D-CNN(All Values in %)

Model	Stress Level	Validation Set				Test Set			
		Acc	**Prec**	**Rec**	**F1**	**Acc**	**Prec**	**Rec**	**F1**
Random Forest	1		77.54	85.51	81.33		78.07	83.18	80.54
	2		60.94	46.99	53.06		73.85	57.83	64.86
	3		68.00	55.74	61.26		64.00	52.46	57.66
	4		53.33	61.54	57.14		44.44	61.54	51.61
	5		82.76	87.27	84.96		80.17	87.39	83.62
	avg	**74.84↓**	**68.51↓**	**67.41↓**	**67.55↓**	**75.31↓**	**68.11↓**	**68.48↓**	**67.66↓**
LSTM	1		87.10	88.32	87.70		84.33	85.51	84.92
	2		82.86	69.88	75.82		75.71	63.86	69.28
	3		68.18	73.77	70.87		72.22	63.93	67.83
	4		50.00	46.15	48.00		63.64	53.85	58.33
	5		87.93	92.73	90.27		80.77	94.59	87.14
	avg	**83.16↑**	**75.21↑**	74.17	**74.53↑**	80.29	75.33	72.35	73.50
1D-CNN	1		92.74	77.57	84.48		89.95	79.44	84.37
	2		79.07	81.93	80.47		77.11	77.11	77.11
	3		68.12	77.05	72.31		68.66	75.41	71.88
	4		30.43	53.85	38.89		64.29	69.23	66.67
	5		82.26	92.73	87.18		83.72	97.30	90.00
	avg	81.08	70.52	**76.62↑**	72.67	**82.37↑**	**76.74↑**	**79.70↑**	**78.00↑**

5.4.2.2 Stress Level Classification: In Vitro

This section compares Random Forest, LSTM, and 1D-CNN models for classifying self-reported stress levels (1–5) in the controlled, in vitro setting, where only physiological and a reduced set of environmental features are available (see Table 5.1). LSTM and 1D-CNN utilise the entire 30-second time series per segment, while Random Forest operates on the last second of each segment as a snapshot-based analysis.

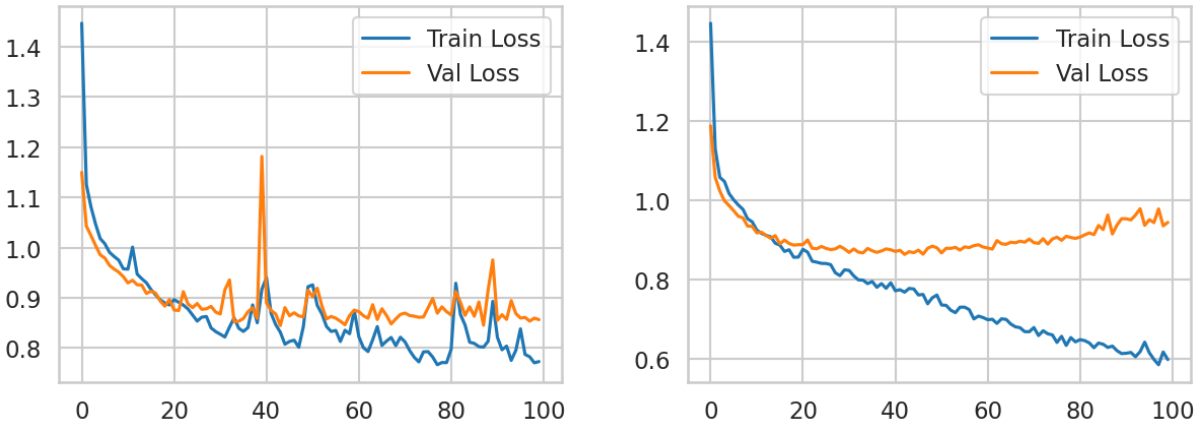


Figure 5.27: Training and validation loss curves for LSTM (left) and 1D-CNN (right) models in self-reported stress level classification (In Vitro).

Figure 5.27 displays the loss curves for both deep learning models. Both LSTM and 1D-CNN show a decreasing trend in training and validation loss, with 1D-CNN demonstrating smoother convergence.

Model accuracy trends for validation and training sets are shown in Figure 5.28. Both models achieve increased accuracy over epochs, and 1D-CNN generally demonstrates greater stability and generalization than LSTM.

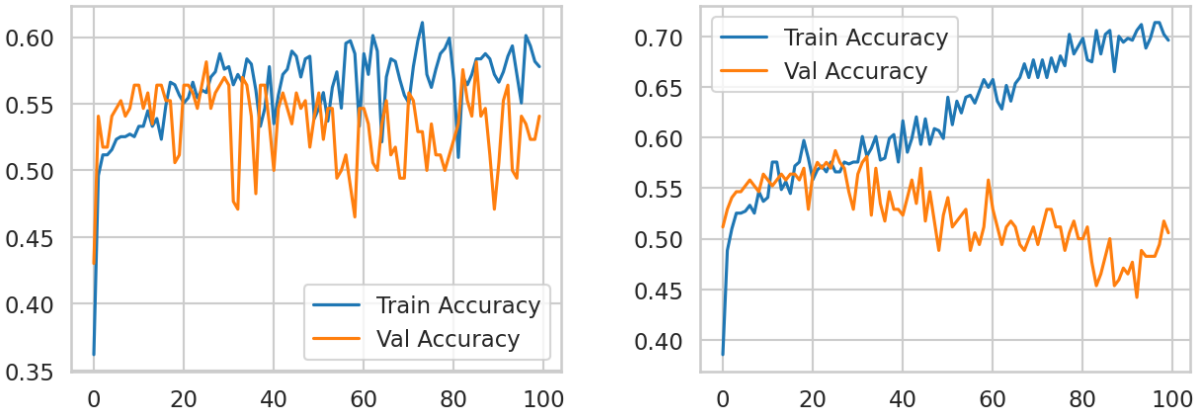


Figure 5.28: Training and validation accuracy curves for LSTM (left) and 1D-CNN (right) models in self-reported stress level classification (In Vitro).

Model performance on the validation set is summarised by the confusion matrices and their normalisations in Figure 5.29 and Figure 5.30: RF (left), LSTM (centre), and 1D-CNN (right). These illustrate the varying abilities of each model to distinguish among stress levels, with the deep models showing relatively higher accuracy at the extreme ends of the scale.

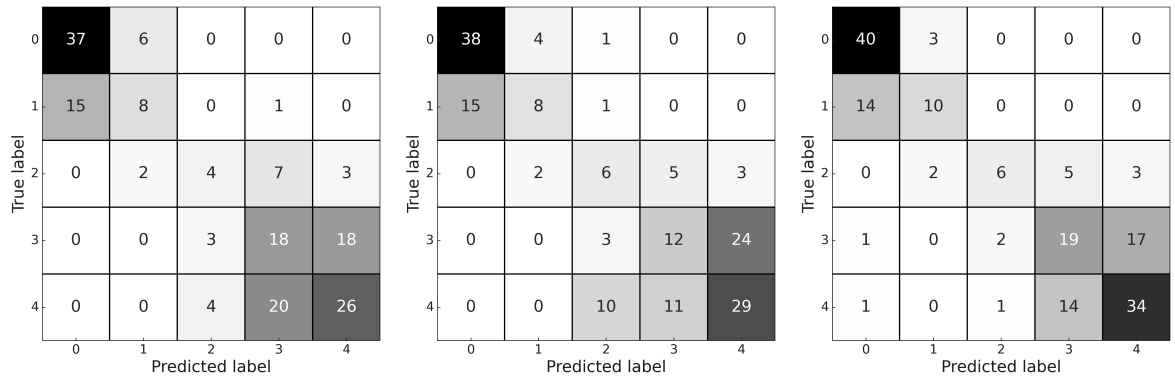


Figure 5.29: Validation confusion matrices for self-reported stress level classification (RF: left, LSTM: centre, 1D-CNN: right) in In Vitro experiments.

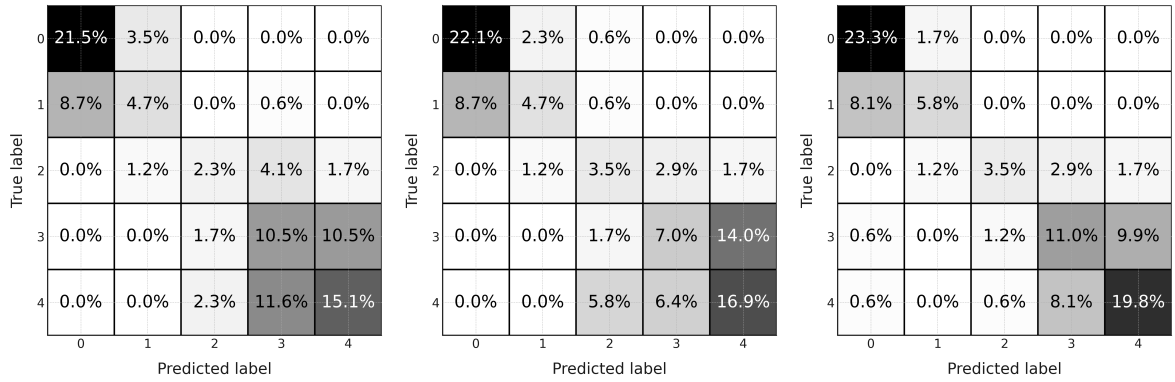


Figure 5.30: Normalised validation confusion matrices for self-reported stress level classification (RF: left, LSTM: centre, 1D-CNN: right) in In Vitro experiments.

Figure 5.31 and 5.32 presents the test set confusion matrices and their normalisations for each model. As in the validation set, the models tend to better identify extreme stress categories, while confusion persists among the mid-range classes, especially for Random Forest and LSTM.

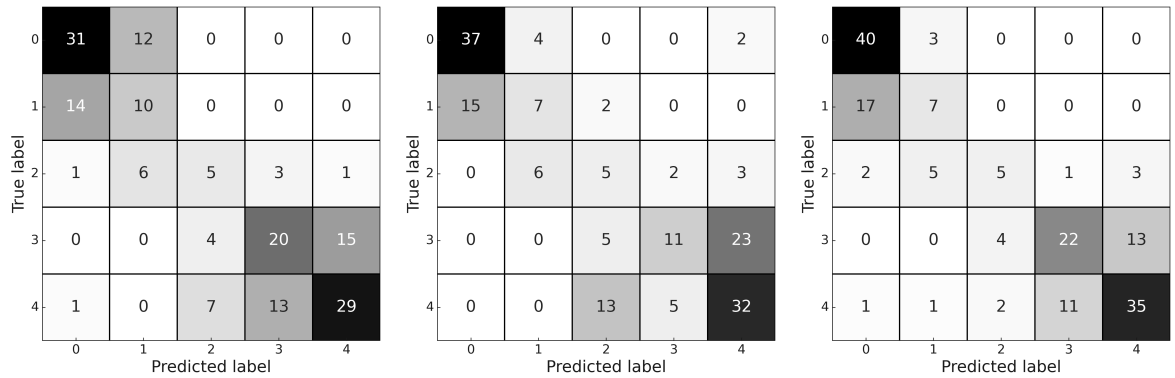


Figure 5.31: Test confusion matrices for self-reported stress level classification (RF: left, LSTM: centre, 1D-CNN: right) in In Vitro experiments.

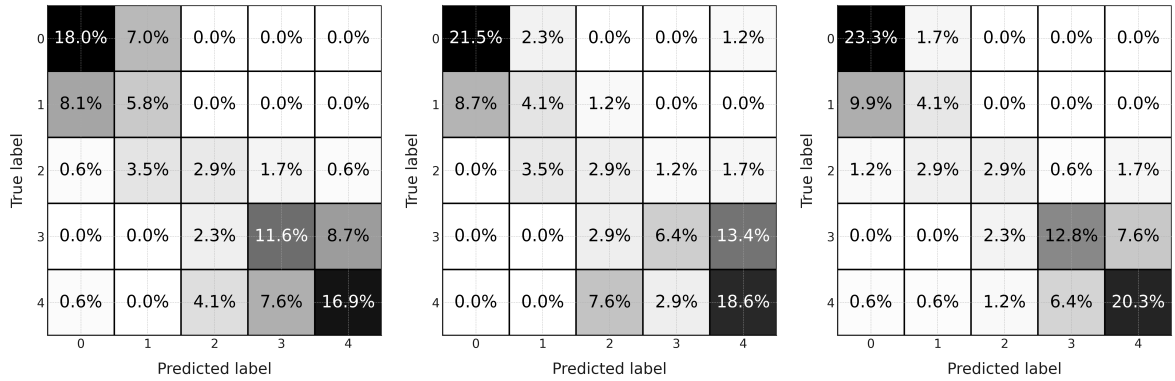


Figure 5.32: Normalised test confusion matrices for self-reported stress level classification (RF: left, LSTM: centre, 1D-CNN: right) in In Vitro experiments.

Random Forest feature importance for In Vitro experiments is illustrated in Figure 5.33. Here, air pressure, wrist temperature, and EDA were among the most impactful predictors for stress level classification, reflecting the strong contribution of these physiological measures in the absence of richer environmental features.

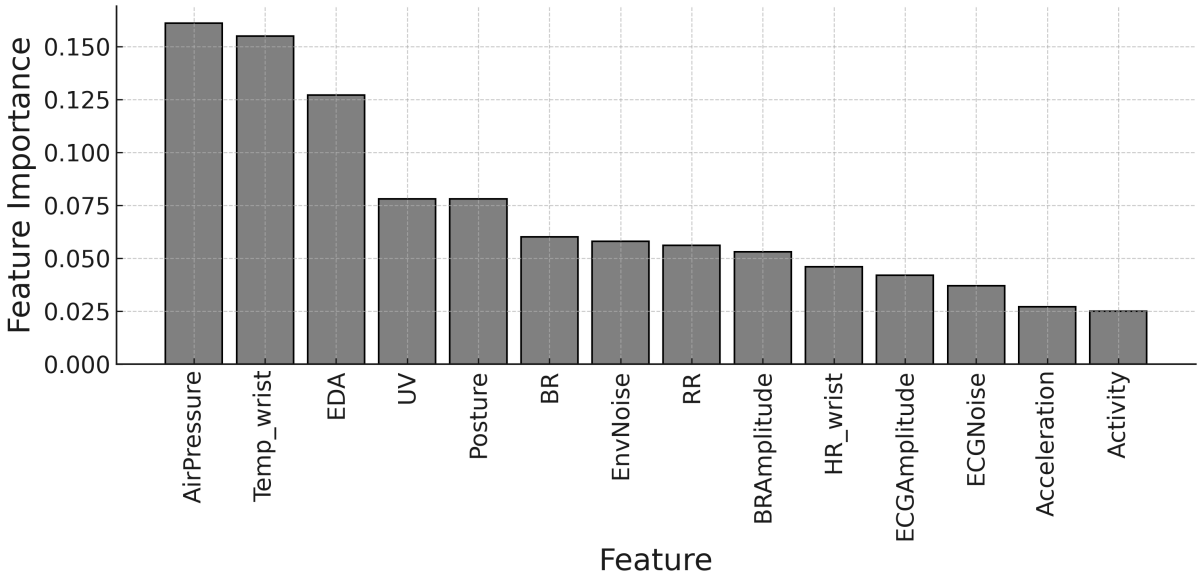


Figure 5.33: Feature importance from Random Forest classifier for self-reported stress level classification in In Vitro experiments.

A summary of model performance for each stress level, with validation and test metrics (accuracy, precision, recall, F1), is presented in Table 5.7. Consistent with the in situ findings, the 1D-CNN model achieved the highest average accuracy, especially for the validation set, confirming its advantage in leveraging sequential patterns even with fewer input features.

Table 5.7: Stress Level Classification (In Vitro): Random Forest vs. LSTM vs. 1D-CNN (All Values in %)

Model	**Stress Level**	**Validation Set**				**Test Set**			
		Acc	**Prec**	**Rec**	**F1**	**Acc**	**Prec**	**Rec**	**F1**
Random Forest	1		71.15	86.05	77.89		65.96	72.09	68.89
	2		50.00	33.33	40.00		35.71	41.67	38.46
	3		36.36	25.00	29.63		31.25	31.25	31.25
	4		39.13	46.15	42.35		55.56	51.28	53.33
	5		55.32	52.00	53.61		64.44	58.00	61.05
	avg	**54.07**↓	**50.39**↓	**48.51**↓	**48.70**↓	55.23	50.58	50.86	50.60
LSTM	1		71.70	88.37	79.17		71.15	86.05	77.89
	2		57.14	33.33	42.11		41.18	29.17	34.15
	3		28.57	37.50	32.43		20.00	31.25	24.39
	4		42.86	30.77	35.82		61.11	28.21	38.60
	5		51.79	58.00	54.72		53.33	64.00	58.18
	avg	**54.07**↓	50.41	49.59	48.85	**53.49**↓	**49.35**↓	**47.73**↓	**46.64**↓
1D-CNN	1		71.43	93.02	80.81		66.67	93.02	77.67
	2		66.67	41.67	51.28		43.75	29.17	35.00
	3		66.67	37.50	48.00		45.45	31.25	37.04
	4		50.00	48.72	49.35		64.71	56.41	60.27
	5		62.96	68.00	65.38		68.63	70.00	69.31
	avg	**63.37**↑	**63.54**↑	**57.78**↑	**58.97**↑	**63.37**↑	**57.84**↑	**55.97**↑	**55.86**↑

5.5 Key Findings

Table 5.8: Summary of Classification Performance for All Models Across Experiments (Test Set Results)

Task	Experiment	Model	Accuracy (%)	F1-Score (%)
Environment Classification	In Situ	Random Forest	—	—
		LSTM	95.23	96.70
		1D-CNN	**97.72**	**98.42**
	In Vitro	Random Forest	—	—
		LSTM	81.40	82.22
		1D-CNN	**94.19**	**94.57**
Stress Level Classification	In Situ	Random Forest	75.31	67.66
		LSTM	80.29	73.50
		1D-CNN	**82.37**	**78.00**
	In Vitro	Random Forest	55.23	50.60
		LSTM	53.49	46.64
		1D-CNN	**63.37**	**55.86**

Table 5.8 summarises the classification performance of all models across each task and experimental setting. For environmental classification, both deep learning approaches outperformed the Random Forest baseline, with 1D-CNN achieving the highest accuracy and F1-score in both in situ and in vitro contexts. Similarly, for self-reported stress level classification, the 1D-CNN model delivered the best results, especially in the in situ scenario with a more comprehensive feature set. These results reinforce the advantage of time-series models for multimodal health and environment analysis.

Random Forest feature importance analysis highlighted the contributions of both physiological (e.g., EDA, RR) and environmental features (e.g., CO2, temperature) in stress prediction. Across all tasks, the models demonstrated strongest performance for extreme stress states, while intermediate states remained more challenging to classify.

Comparison with Prior Work: These findings are also significant when compared with

existing literature, particularly the *DigitalExposome* study by Johnson et al. [173], which aimed to understand environmental impacts on wellbeing using a Deep Belief Network (DBN) and a smaller, custom dataset. In contrast, my research leverages larger, publicly available datasets with a more diverse participant base, enabling greater generalizability. Additionally, my methodological approach employs state-of-the-art deep learning models (LSTM, 1D-CNN) with systematic hyperparameter tuning and cross-validation, achieving superior predictive performance. Unlike *DigitalExposome*, which remains primarily conceptual, this study offers practical applications in healthcare diagnostics and urban analytics by directly modelling physiological stress and environmental conditions with high resolution and interpretability.

These findings support the utility of multimodal, sequential data and advanced deep learning techniques for real-world wellbeing and stress monitoring. The following chapter details a dedicated framework for physiological state prediction using wearable sensor data.

Chapter 6

Machine Learning and Deep Learning Framework for Predicting Physiological States

6.1 Introduction

Classifying whether someone is relaxed, neutral or in a stressed state from physiological markers has multiple real world uses. For example, carers of patients, especially those suffering from mental health issues, could be alerted if the patient is flagged as being in a high stress state so that the carer can prevent any harm to the patient.

Another use-case is self improvement, individuals could track their stress level throughout the day and aim to minimise stress throughout the day using various methods such as breathing techniques, social activities, etc. Consistently high levels of stress have been shown to cause a large array of diseases and disorders [189], so trying to minimise levels of stress would be beneficial. The first step to doing so would be to track stress to monitor progress. To this end, the EnviroWellBeing dataset from this work can be used alongside a number of ML and DL algorithms.

Although a similar problem to the one proposed in this chapter was tackled within chapter 5.

In that chapter, the stress label was a discrete integer in 1, 2, 3, 4, 5. Naturally, for practical applications, it makes more sense to further discretise these labels into three labels; relaxed, neutral and stressed. Furthermore, in this chapter, only physiological data from the wristband is used, as opposed to the use of both the wristband and the chestbelt. Again, this is due to practical considerations, users of a stress prediction application should not be expected to wear a chestbelt at all times as it can be quite uncomfortable to do so. However, most users would be happy with wearing the wristband solely for continuous stress prediction throughout everyday life.

As stated in prior chapters, the dataset was collected using 53 participants. Each participant was asked to walk a fixed distance in a polluted and unpolluted area. The polluted area was on the sidewalk of a busy road and the unpolluted area was in a quiet park. During the walk, participants wore a wrist band that tracks the following physiological markers at a frequency of 8 Hz:

1. Electrodermal Activity

2. Heart Rate

3. Ultraviolet Light Levels

4. Noise levels

5. Air Pressure

6. Body Temperature

They also wore a chestbelt and data-logger which track extra physiological markers and environment pollution levels. The data from the chest-belt and data-logger were not used for the task of realistically classifying stress states, as in a real life inference setting, users would at most be able to wear the wrist band, as it would be too inconvenient and uncomfortable to also be constantly wearing the chest-belt and data-logger.

The walks lasted an average of 7 minutes and at any time participants could indicate their stress level from 1 to 5. The experiments were repeated using a simulated environment. For the simulated polluted environment, each participant was placed in an controlled room setting

alone, where annoying sounds such as engines revving and car horns were played. This lasted for 2 minutes. For the simulated unpolluted environment, the same procedure was followed, except that pleasant sounds were played such as birds whistling and the light breeze of the wind, for a further granular description about data collection process please refer to section 5.3.

6.2 Data Preprocessing

The dataset consisted of 212 experiments, resulting from 53 participants completing four different scenarios each. All experiments both simulated and real-world were included, as they provided valuable information for predicting stress. Initially, the physiological data, sampled at 8 Hz, was downsampled to 1 Hz. Exploratory analysis showed that the physiological markers did not fluctuate significantly within a single second, making 1 Hz sufficient for capturing relevant changes while reducing data volume.

Given the relatively small number of samples compared to the number of features and time-steps, the dimensionality of the data presented a major challenge. High-dimensional datasets tend to suffer from sparsity, making it difficult for models to extract meaningful patterns. As the feature space grows exponentially, so does the risk of overfitting. Additionally, processing such data is computationally expensive, with many algorithms scaling poorly in both time and resource requirements as dimensionality increases. In time-series data specifically, every additional time-step increases the input space exponentially, often without adding meaningful predictive value. Therefore, reducing the number of time-steps was a necessary strategy to simplify the model space and improve generalisation performance.

Each experiment was then split into smaller segments, or sub-experiments, each one minute long comprising 60 time-steps at 1 Hz. Each segment was assigned a single label based on the average of the self-reported stress scores during that minute. If the average fell within the range [1, 2.5], the label was set to ”relaxed”; between (2.5, 3.5), the label was ”neutral”; and between [3.5, 5], it was ”stressed.” This segmentation offered two key advantages: First, it helped reduce overfitting by limiting the number of time-steps per sample to the most relevant window of recent physiological data. Second, it enabled the use of a broader range of models,

including both sequence-based architectures like LSTMs and fixed-shape models like MLPs, since all data points had consistent dimensions.

Finally, the dataset was standardized to ensure all features had zero mean and unit variance. The resulting dataset contained 493 samples, each represented as a (60, 6) matrix corresponding to 60 time-steps of six physiological features, along with a label indicating one of three stress levels.

6.3 Model Configuration and Training

The dataset was divided into training, validation, and test subsets, with proportions of 60%, 20%, and 20%, respectively. The validation set was specifically used for hyperparameter tuning to ensure optimal model performance during training. Several models were evaluated as part of this process. For a more detailed explanation of the machine learning modules used, please refer to section 3.4. the following are the Architecture Summary for DTW-KNN, DTW-SVM, LSTM, 1D-CNN, and Multi-Layer Perceptron (MLP) Models.

The first model evaluated was DTW-KNN, a k-nearest neighbours approach using DTW as the distance metric. A value of K = 1 was selected. DTW is widely recognised as a leading distance metric for time-series data [169], which motivated its use in this context.

The second model was DTW-SVM, which applies Support Vector Machines using a DTW kernel to handle sequential inputs effectively.

The third model evaluated is 1D Convolutional Neural Network (1D-CNN). The 1D-CNN model used for stress level classification is designed to effectively capture temporal patterns in the physiological time series data.

The input shape for the 1D-CNN is (batch_size, sequence_length, features). The convolutional layers include 3 layers with the following configuration: Layer 1 has 64 filters with kernel size 5, Layer 2 has 128 filters with a kernel size 5, and Layer 3 has 128 filters with a kernel size 3. After each convolutional layer, a ReLU activation function is applied, followed by max pooling

with kernel size of 2 and a stride 2. Global adaptive average pooling is used to reduce the sequence dimension.

The fully connected layers consist of two parts: FC1 transforms 128 units to 64 units with ReLU activation, and FC2 transforms 64 units to 3 units (output layer). A dropout rate of 0.2 is applied after global pooling and between fully connected layers. The output represents 3 classes: Relaxed, Neutral, and Stressed.

For training, a batch size of 8 is used with the Adam optimiser and Cross-Entropy Loss function. Gradient clipping is set at 5.0, and early stopping is implemented with a patience of 100 epochs and a minimum delta of 0. The maximum number of epochs is set to 10,000, though early stopping typically prevents reaching this limit.

The 1D-CNN architecture transforms the input time series data by applying convolutional operations that can detect local patterns in the data, regardless of where they occur in the time series. The model gradually increases the number of filters (from 64 to 128) to extract increasingly complex features while reducing the temporal dimension through max pooling. The global average pooling layer further reduces dimensionality before classification through fully connected layers.

The fourth model used is the LSTM, which is designed to capture long-range dependencies and temporal dynamics in the physiological time series data for stress level classification.

The input shape for this model is also (batch_size, sequence_length, features). The LSTM configuration includes a hidden size of 64 units with 2 layers. The model is bidirectional, meaning it reads the sequence both forward and backward, and has a dropout rate of 0.2 between LSTM layers. For output processing, the model uses the final time step output.

The fully connected layers include FC1, which transforms 128 units (2×64 due to bidirectionality) to 64 units with ReLU activation, and FC2, which transforms 64 units to 3 units (output layer). A dropout rate of 0.2 is applied between fully connected layers. Like the CNN model, the output represents 3 classes: Relaxed, Neutral, and Stressed.

Training parameters are similar to the CNN model: batch size of 8, Adam optimizer, Cross-

Entropy Loss, gradient clipping at 5.0, early stopping with patience of 100 epochs and minimum delta of 0, and a maximum of 10,000 epochs.

The LSTM architecture processes the time series sequentially, maintaining internal memory states that allow it to learn dependencies across different time steps. The bidirectional approach enables the model to incorporate information from both past and future time steps, which can be particularly valuable for stress detection where context from the entire sequence may be important. The final time step output captures the accumulated information from the entire sequence, which is then passed through fully connected layers for classification.

The fifth model evaluated is the MLP, this model serves as a baseline approach for stress level classification using the physiological time series data. Unlike the CNN and LSTM models that preserve the temporal structure of the data, the MLP flattens the entire time series input before processing.

The input shape for this model is also (batch_size, sequence_length, features), but the first operation flattens this to (batch_size, sequence_length * features). The MLP architecture consists of two hidden layers with sizes [32, 16], considerably smaller than those used in the CNN and LSTM models. Each hidden layer is followed by a ReLU activation function. A dropout rate of 0.1 is applied between all layers to prevent overfitting, slightly lower than the 0.2 used in the other models.

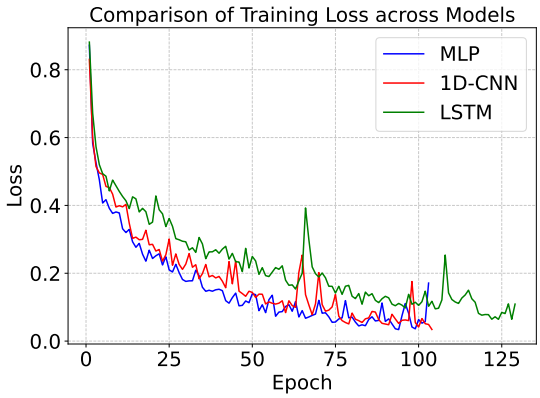
The training parameters for the MLP are similar to those of the other models: batch size of 8, Adam optimizer, Cross-Entropy Loss, gradient clipping at 5.0, early stopping with patience of 100 epochs and minimum delta of 0, and a maximum of 10,000 epochs. Like the other models, the output layer has 3 units corresponding to the three stress levels: Relaxed, Neutral, and Stressed.

The MLP approach, while simpler than CNN and LSTM, provides an important benchmark for comparison. Its architecture does not explicitly model temporal dependencies or local patterns in the sequence but instead treats the entire flattened sequence as a single feature vector. This makes it useful for evaluating whether the more complex temporal modelling approaches of CNN and LSTM provide meaningful improvements for stress classification based on physio-

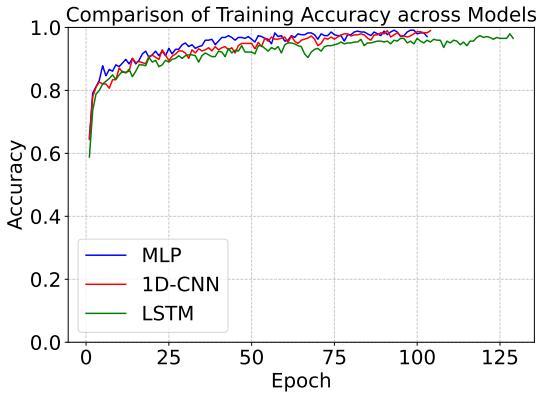
logical data.

All three models (1D-CNN, LSTM, MLP) were trained using a 60-20-20 split for training, validation, and testing, with early stopping based on validation loss to prevent overfitting. The training metrics (accuracy and loss) for both training and validation sets were tracked using TensorBoard, allowing for visual comparison of model performance over time.

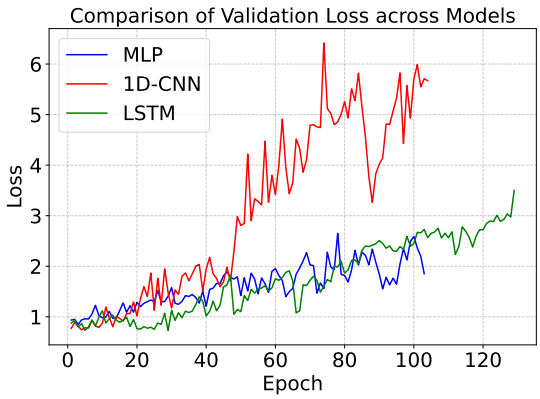
Figure 6.1 shows a comparative plot of model performance for 1D-CNN, LSTM and MLP training curves.



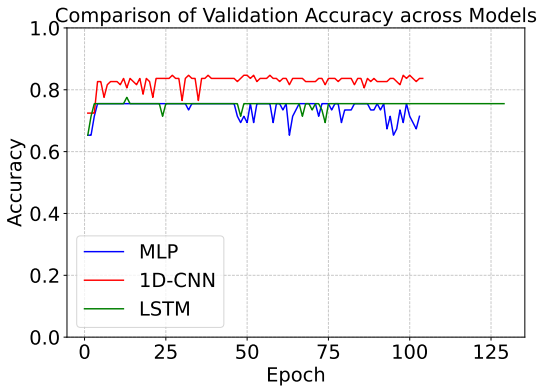
(a) Comparative training loss.



(b) Comparative training accuracy.



(c) Comparative validation loss.



(d) Comparative validation accuracy.

Figure 6.1: A comparative plot of model performance for 1D-CNN, LSTM and MLP training curves

When comparing the validation loss across the three models MLP, 1D-CNN, and LSTM distinct patterns of overfitting can be observed. Both the MLP and 1D-CNN models exhibit signs of overfitting starting from epoch 10. Although the training loss for these models continues to decrease steadily, indicating that the models are learning the training data well, the validation

loss begins to increase after epoch 10. This divergence suggests that the models are no longer generalizing effectively to unseen data and are instead memorizing the training data, a classic sign of overfitting.

In contrast, the LSTM model demonstrates better generalization for a longer period, with validation loss remaining stable until epoch 25. After epoch 25, however, the validation loss also begins to rise while the training loss continues to decline, indicating that overfitting starts at that point. Among the three models, the LSTM exhibits the most robust performance in terms of generalization, as it maintains a stable validation loss for more epochs before overfitting begins.

6.4 Results and Discussion

The models DTW-KNN, DTW-SVM, MLP, 1D-CNN and LSTM were evaluated on the held-out test set, which comprised 99 data points. The class distribution in the test set was: 46 samples labelled as Stressed (Class 0), 15 as Neutral (Class 1), and 38 as Relaxed (Class 2).

Table 6.1 presents the overall test set accuracy and macro F1-score for the evaluated models.

Table 6.1: Results for classifying relaxed, neutral and stressed states on the test set.

Model	Accuracy (%)	F1 (Macro Avg)
DTW-KNN	**73**	**0.68**
LSTM	70	0.5
1D-CNN	68	0.53
DTW-SVM	68	0.62
MLP	61	0.44

Table 6.2 provides a detailed classification report for the best performing model, the KNN-DTW, showing precision, recall, and F1-score for each of the three stress states.

As summarised in Table 6.1, the DTW-KNN algorithm achieved the highest overall accuracy at 73% with a macro-average F1-score of 0.68, followed by LSTM network at 70% accuracy (F1-score 0.50), 1D-CNN at 68% accuracy (F1-score 0.53), DTW-SVM at 68% accuracy (F1-score 0.62) and MLP at 61% accuracy (F1-score 0.44). This ranking suggests that for this particular

Table 6.2: Summarising performance metrics precision, recall, F1-score across mental states for the best performing model KNN-DTW.

Class	Precision	Recall	F1-Score	Support
Stressed (Class 0)	0.76	0.84	0.80	38
Neutral (Class 1)	0.50	0.07	0.12	15
Relaxed (Class 2)	0.71	0.85	0.77	38
Accuracy			0.73	99
Macro Avg	0.66	0.59	0.56	99
Weighted Avg	0.70	0.73	0.68	99

stress classification task, traditional methods like DTW-KNN performed competitively or even better than more complex deep learning approaches.

When examining class-specific performance, as detailed in Table 6.2, the DTW-KNN model demonstrated the highest overall accuracy. For the ”Stressed” class (Class 0), the model achieved precision 0.76, recall 0.84, and F1-score 0.80. For the ”Neutral” class (Class 1), it showed precision 0.50, recall 0.07, and F1-score 0.12. For the ”Relaxed” class (Class 2), it demonstrated precision 0.71, recall 0.85, and F1-score 0.77. Overall, it achieved balanced performance across both ”Relaxed” and ”Stressed” classes with high F1-scores, resulting in a macro average precision of 0.66, recall of 0.59, and F1-score of 0.56. This suggests that for this dataset, local patterns in the feature space captured by the DTW-KNN approach are highly informative for stress classification.

The LSTM network achieved the second-highest accuracy (70%) and demonstrated particularly strong performance in identifying the ”Relaxed” class (recall 0.91). Its balanced F1-scores for both ”Relaxed” and ”Stressed” classes (0.76 and 0.75) indicate effective learning of temporal patterns in the physiological data.

The 1D-CNN model achieved a a moderate efficacy with an accuracy of 68% across the labels for the three physical states. The model exhibited strong discriminative capabilities for both the Relaxed class (precision=0.69, recall=0.79, f1-score=0.73) and the Stressed class (precision=0.71, recall=0.79, f1-score=0.75), indicating reliable performance in detecting these emo-

tional states from physiological time series data. Moreover, The architecture, which comprises three convolutional layers with increasing filter complexity (64, 128, 128 filters) and decreasing kernel sizes (5, 5, 3), effectively captured temporal patterns in the physiological data.

The DTW-SVM model achieved an accuracy of 68% with good performance on both ”Relaxed” and ”Stressed” classes (F1-scores of 0.72 and 0.74). The model showed a particular strength in recall for the ”Stressed” class (0.84), suggesting effectiveness in identifying stress states.

The MLP achieved a moderate accuracy of 61% with balanced performance across ”Relaxed” and ”Stressed” classes (F1-scores of 0.65 and 0.67). Despite its architectural simplicity compared to other neural networks, it performed reasonably well, suggesting that even without explicit modelling of temporal relationships, the flattened time series data contains distinguishable patterns for stress classification.

A consistent finding across all models was the difficulty in classifying the ”Neutral” stress state. As shown in Table 6.2, only the DTW-KNN model showed any ability to identify this class, and even then with very low recall (0.07). This suggests that the physiological signals for the intermediate stress level may not be sufficiently distinct from the ”Relaxed” and ”Stressed” states in the feature space, or that this class may be underrepresented in the training data.

The comparative performance analysis of these five models, as summarised in Table 6.1, provides several insights into the task of stress level classification from physiological time series data.

6.5 Key Findings

In conclusion, developers of stress prediction applications can use the stress prediction framework outlined here along with the EnviroWellBeing dataset to obtain a reliable stress prediction model. Results could be further improved by testing other models such as Random Forest.

Furthermore, expanding the EnviroWellBeing dataset to include more participants is likely to

decrease the generalisation error of the models, making them more accurate predictors of stress. The EnviroWellBeing dataset could be expanded, not only by including more participants, but also including further environments on top of the peaceful simulated, noisy simulated, green and polluted environments. Below is a list of example environments that could be added onto the EnviroWellBeing dataset: a public speaking environment, where participants could be asked to perform a public speech in front of a large crowd, while wearing the wristband to record physiological data and self reporting stress at regular intervals; an exercise environment, where participants could be asked to jog lightly while wearing the wristband to record physiological data and self reporting stress at regular intervals; and a TV environment, where participants could be asked to watch a TV show which they enjoy while wearing the wristband to record physiological data and self reporting stress at regular intervals.

All of this additional data within different environments, especially environments that are encountered within everyday life, will decrease the generalization error of the models and increase their prediction accuracy. Future works should initially focus on those two aspects.

Building upon the comprehensive analysis of environmental and physiological data presented in the preceding chapters, Chapter 7 revisits the core research questions to critically evaluate how effectively the study's methodologies and findings address them. It synthesises the major contributions such as the development of the EnviroWellBeing Dataset and the implementation of machine learning models for stress and mental health prediction while highlighting their broader implications for environmental health research. This concluding chapter ties together the empirical insights and methodological innovations to reflect on their significance, laying the groundwork for future research in the dynamic interface between environmental conditions and human wellbeing.

Chapter 7

Conclusion and Future Work

This chapter synthesises the principal contributions and findings of the thesis, positioning them within the broader landscape of environmental health research and sensor-driven wellbeing analytics. The research commenced with the objective of systematically investigating the dynamic relationship between environmental exposures and human physiological states, leveraging advanced sensing technologies and machine learning methodologies.

7.1 Addressing Research Questions

Throughout the preceding chapters, the core research questions (RQ1–RQ4) guided the study design, data collection, and analytical approaches. The implementation of a robust multimodal data collection methodology (Chapter 4) enabled the creation of the EnviroWellBeing Dataset, directly addressing RQ1 regarding practical integration of biometric and environmental sensing in both real-world (*in situ*) and controlled (*in vitro*) settings.

Subsequent analysis (Chapter 5) demonstrated that deep learning models, particularly 1D-CNN, can accurately differentiate between environmental conditions based on physiological signals alone, addressing RQ2. These findings indicate that specific environmental exposures elicit distinct physiological responses that are detectable by advanced models.

Expanding on this, integrated modelling of environmental and physiological data enabled robust

prediction of self-reported stress levels (RQ3), with model performance further improved by richer feature sets. Random Forest feature importance analysis highlighted both environmental and physiological predictors. The value of temporal sequence modelling (RQ4) was confirmed, as sequential models leveraging time-series data outperformed snapshot-based approaches in capturing stress dynamics.

The following section summarises the main empirical results from each chapter, highlighting how the research questions were answered in practice.

7.2 Summary of Key Results

A summary of major findings is presented here, mapped to the progression of the thesis chapters.

In Chapter 3, classical machine learning algorithms (logistic regression, support vector machines, random forest) were shown to effectively distinguish depressed patients from healthy controls using wearable sensor data, with the best model achieving a test accuracy of **83.4%**.

The creation and deployment of the EnviroWellBeing Dataset (Chapter 4) established a comprehensive resource for integrating multimodal physiological and environmental data in both laboratory and real-world contexts.

Chapter 5 advanced the analysis to deep learning approaches for environmental classification and stress prediction. The 1D-CNN model achieved test accuracies of **97.7%** (in situ) and **94.2%** (in vitro) for environmental classification, and **82.4%** (in situ) and **63.4%** (in vitro) for self-reported stress level prediction. Both environmental and physiological features contributed significantly to these outcomes.

Finally, Chapter 6 presented a practical framework for stress state classification using only wristband data, with the DTW-KNN model reaching a test set accuracy of **73.0%** for classifying relaxed, neutral, and stressed states.

Together, these results demonstrate the feasibility and value of combining multimodal sensor data with advanced machine learning techniques for objective wellbeing assessment.

7.3 Concluding Remarks

The synthesis of methods and results across Chapters 1 through 6 demonstrates a coherent trajectory from motivation and literature review to methodological innovation and empirical insight. The research motivation and significance were established in Chapter 1, while Chapter 2 highlighted critical gaps in the literature—particularly the need for comprehensive datasets and integrated data analytics. The subsequent chapters addressed these gaps through comparative analysis, novel dataset construction, and rigorous model evaluation.

By integrating findings from all chapters, this thesis underscores the importance of robust data collection, interdisciplinary methods, and advanced analytics in environmental health research. The empirical advances presented here form a solid foundation for both theoretical understanding and practical applications in wellbeing analytics.

7.4 Limitations

Despite the comprehensive scope of this thesis, several limitations must be acknowledged. Data collection challenges, including sensor malfunctions and session truncation, led to reduced usable data and potential bias in the analysed cohorts (see Chapter 4). The lack of uniformity in session durations, especially in *in situ* experiments, introduced imbalance and variability in the dataset. Technical issues with GPS logging prevented detailed spatial analysis of environmental influences.

Physiological measurements were limited to signals available from wearable sensors; direct heart rate variability was not recorded but inferred from RR intervals, which may introduce detection errors. The cohort was relatively homogeneous in demographic characteristics and environmental exposures, which limits generalisability. Sample sizes, further reduced by segmentation and data cleaning, restricted statistical power for some analyses.

From a modelling perspective, deep learning models such as LSTM and 1D-CNN (Chapter 5) present interpretability challenges and require substantial computational resources. The stress prediction framework developed in Chapter 6 was evaluated on a limited dataset; its generalis-

ability to new or more diverse populations remains untested, and model performance for neutral stress classification was suboptimal.

Additionally, while the thesis aimed to explore the practical deployment of the proposed framework, the development of a real-time mobile or wearable application could not be pursued due to technical constraints. Specifically, the Microsoft Band device used for data collection was discontinued, and no compatible or officially supported libraries exist to enable real-time data access on modern platforms. This limitation prevented the implementation and validation of a mobile app during the study. Future research should consider using more current wearable devices with active developer support to overcome these technical barriers and facilitate mobile deployment.

7.5 Future Work

Future research should address these limitations by expanding data collection to include more diverse participant groups and longer-term, longitudinal designs, improving generalisability and enabling the study of cumulative effects. Enriching the dataset with additional physiological markers, direct HRV measurement, and higher-resolution environmental sensors will provide greater analytic depth.

Advances in modelling should prioritise interpretability such as the application of explainable AI methods and efficiency, to enable real-time inference on wearable devices. Further validation of the stress prediction framework in ambulatory, real-world settings is necessary, with particular focus on improving neutral state detection and adapting models for broader deployment.

Collaboration with researchers in environmental science, health, and engineering should facilitate richer contextual annotation, innovative analysis strategies, and the translation of findings into practical interventions. Finally, public release of anonymised datasets and code, following ethical guidelines, will support transparency, replication, and future advances in environmental health analytics.

Development of a mobile or wearable application for real-time stress prediction should also be pursued in future work. Such an application would not only enhance the framework's practical

utility but also enable personalised feedback and continuous monitoring in naturalistic environments. Future studies should include testing these mobile solutions in clinical or controlled settings to evaluate their real-world impact. While this direction could not be fully explored within the current study due to technical constraints specifically, the discontinuation of the Microsoft Band and the lack of compatible, officially supported libraries for modern platforms these limitations have been clearly identified. The discontinuation of the Microsoft Band product line and absence of compatible APIs prevented real-time data access and thus hindered mobile deployment in this research. Future work should therefore employ up-to-date wearable devices that support robust data acquisition and mobile integration, enabling broader applicability and experimental validation of the proposed framework in real-world and clinical contexts

Bibliography

[1] Y. Trossman-Haifler and D. Fisher-Gewirtzman, "How urban wellbeing is influenced by spatial urban parameters (density, morphology, vegetation & commerce), as examined in a vr framework," *Architectural Science Review*, vol. 65, no. 5, pp. 370–384, 2022.

[2] C. Zhang, J. Li, T. Liu, M. Xu, H. Wang, and X. Li, "The spatiotemporal evolution and influencing factors of the chinese cities' ecological welfare performance," *International Journal of Environmental Research and Public Health*, vol. 19, no. 19, p. 12955, 2022.

[3] G. Gizachew, "Spatial-temporal and factors influencing the distribution of biodiversity: A review," *ASEAN Journal of Science and Engineering*, 2022.

[4] P. Lignier, D. Jarvis, D. Grainger, and T. Chaiechi, "Does the climate impact satisfaction with life? an australian spatial study," *Weather, Climate, and Society*, vol. 15, no. 1, pp. 159–175, 2023.

[5] L. Sun, G. Zhang, D. Zhao, L. Ji, H. Gu, L. Sun, and X. Li, "Explore the correlation between environmental factors and the spatial distribution of property crime," *ISPRS International Journal of Geo-Information*, vol. 11, no. 8, p. 428, 2022.

[6] P. Lauriola, H. Crabbe, B. Behbod, F. Yip, S. Medina, J. C. Semenza, S. Vardoulakis, D. Kass, A. Zeka, I. Khonelidze *et al.*, "Advancing global health through environmental and public health tracking," *International journal of environmental research and public health*, vol. 17, no. 6, p. 1976, 2020.

[7] J. Arebola, P. Martin-Olmedo, F. Pérez-Carrascosa, J. León, R. Echeverría, I. Salcedo-Bellido, and C. Gomez-Peña, "Integrated approaches for the assessment of health im-

pacts of environmental chemicals: our experience in the gramo cohort study," *European Journal of Public Health*, vol. 30, no. Supplement_5, pp. ckaa165–149, 2020.

[8] C. Sampaio, E. Macedo, M. C. Coelho, and J. M. Bandeira, "Economic and environmental analysis of measures from a sustainability urban mobility plan–application to a small sized city," *Transportation Research Procedia*, vol. 48, pp. 2580–2588, 2020.

[9] B. Zhang, S. Su, Y. Zhu, and X. Li, "An lca-based environmental impact assessment model for regulatory planning," *Environmental Impact Assessment Review*, vol. 83, p. 106406, 2020.

[10] K. Movahed and H. Nikounam Nezami, "Evaluation and prioritization of indicators that improve the quality of life in residential neighborhoods," *International journal of architecture and urban development*, vol. 12, no. 2, pp. 59–72, 2022.

[11] F. Sharmin, "Towards the valuation of open spaces: a hedonic based investigation for sustainable planning in the dense urban context of dhaka," *American Journal of Environmental and Resource Economics*, vol. 5, no. 4, pp. 97–103, 2020.

[12] J. M. Robinson and M. F. Breed, "The lovebug effect: Is the human biophilic drive influenced by interactions between the host, the environment, and the microbiome?" *Science of the Total Environment*, vol. 720, p. 137626, 2020.

[13] L. Martin, M. P. White, A. Hunt, M. Richardson, S. Pahl, and J. Burt, "Nature contact, nature connectedness and associations with health, wellbeing and pro-environmental behaviours," *Journal of environmental psychology*, vol. 68, p. 101389, 2020.

[14] I. S. Bower, G. M. Clark, R. Tucker, A. T. Hill, J. A. Lum, M. A. Mortimer, and P. G. Enticott, "Enlarged interior built environment scale modulates high-frequency eeg oscillations," *Eneuro*, vol. 9, no. 5, 2022.

[15] M. J. Sirgy, "Effects of biology, drugs, life events, and the environment on wellbeing," in *The Psychology of Quality of Life: Wellbeing and Positive Mental Health*. Springer, 2021, pp. 175–204.

[16] S. J. Zawadzki, L. Steg, and T. Bouman, "Meta-analytic evidence for a robust and positive association between individuals' pro-environmental behaviors and their subjective wellbeing," *Environmental Research Letters*, vol. 15, no. 12, p. 123007, 2020.

[17] A. R. Olubi and H. A. Ayoola, "The spatial impacts of covid-19 and households behavioural changes in akure, nigeria," *European Journal of Humanities and Social Sciences*, vol. 1, no. 6, pp. 113–121, 2021.

[18] D. Buil-Gil and S. Langton, "Gis and geovisual analysis," *Sage research methods*, 2020.

[19] P. J. B. R. A. Davis, *Introduction to time series and forecasting*. springer publication, 2016.

[20] C. Lv, C. Shao, and C.-C. Lee, "Green technology innovation and financial development: Do environmental regulation and innovation output matter?" *Energy Economics*, vol. 98, p. 105237, 2021.

[21] M. F. Bashir, B. Ma, M. Shahbaz, and Z. Jiao, "The nexus between environmental tax and carbon emissions with the roles of environmental technology and financial development," *Plos one*, vol. 15, no. 11, p. e0242412, 2020.

[22] A. E. Caglar, M. Mert, and G. Boluk, "Testing the role of information and communication technologies and renewable energy consumption in ecological footprint quality: Evidence from world top 10 pollutant footprint countries," *Journal of Cleaner Production*, vol. 298, p. 126784, 2021.

[23] J. Wu, "Research on the impact of technological innovation on environmental pollution–based on the moderating effect of internet development," in *E3S Web of Conferences*, vol. 143. EDP Sciences, 2020, p. 02054.

[24] G. Yang, D. Zha, X. Wang, and Q. Chen, "Exploring the nonlinear association between environmental regulation and carbon intensity in china: The mediating effect of green technology," *Ecological Indicators*, vol. 114, p. 106309, 2020.

[25] A. Moosa and F. He, "The relationship between green operation and sustainable quality performance: The mediation role of environmental technology," *Journal of Environmen-*

tal Planning and Management, vol. 65, no. 8, pp. 1414–1435, 2022.

[26] Y. Gao, L. Liu, B. Hu, T. Lei, and H. Ma, "Federated region-learning for environment sensing in edge computing system," *IEEE Transactions on Network Science and Engineering*, vol. 7, no. 4, pp. 2192–2204, 2020.

[27] N. Hassan and C. Woo, "Machine learning application in water quality using satellite data," in *IOP Conference Series: Earth and Environmental Science*. IOP Publishing, 2021, p. 012018.

[28] A. Hamrani, A. Akbarzadeh, and C. A. Madramootoo, "Machine learning for predicting greenhouse gas emissions from agricultural soils," *Science of The Total Environment*, vol. 741, p. 140338, 2020.

[29] F. Serdio, E. Lughofer, K. Pichler, T. Buchegger, M. Pichler, and H. Efendic, "Fault detection in multi-sensor networks based on multivariate time-series models and orthogonal transformations," *Information Fusion*, vol. 20, pp. 272–291, 2014.

[30] T. Schaul, D. Horgan, K. Gregor, and D. Silver, "Universal value function approximators," in *Proceedings of the 32nd International Conference on Machine Learning*, ser. Proceedings of Machine Learning Research, F. Bach and D. Blei, Eds., vol. 37. Lille, France: PMLR, 07–09 Jul 2015, pp. 1312–1320. [Online]. Available: https://proceedings.mlr.press/v37/schaul15.html

[31] E. Kanjo, "Noisespy: A real-time mobile phone platform for urban noise monitoring and mapping," *Mobile Networks and Applications*, vol. 15, pp. 562–574, 2010.

[32] F. Guerrache, A. Aldabbagh, and E. Kanjo, "Multiple sensor fusion approach to map environmental noise impact on health," in *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct*, 2016, pp. 1074–1078.

[33] F. Guerrache, E. M. Younis, and E. Kanjo, "Quantifying environmental noise impact on heart rate variability." in *MobiCASE*, 2016, pp. 149–150.

[34] Z. Ballard, C. Brown, A. M. Madni, and A. Ozcan, "Machine learning and computation-enabled intelligent sensor design," *Nature Machine Intelligence*, vol. 3, no. 7, pp. 556–565, 2021.

[35] B. Zhang, H. Zhang, G. Zhao, and J. Lian, "Constructing a pm2. 5 concentration prediction model by combining auto-encoder with bi-lstm neural networks," *Environmental Modelling & Software*, vol. 124, p. 104600, 2020.

[36] Y. Sun, X. Wang, N. Ren, Y. Liu, and S. You, "Improved machine learning models by data processing for predicting life-cycle environmental impacts of chemicals," *Environmental Science & Technology*, vol. 57, no. 8, pp. 3434–3444, 2022.

[37] A. A. Dharmasaputro, N. M. Fauzan, M. Kallista, I. P. D. Wibawa, and P. D. Kusuma, "Handling missing and imbalanced data to improve generalization performance of machine learning classifier," in *2021 International Seminar on Machine Learning, Optimization, and Data Science (ISMODE)*. IEEE, 2022, pp. 140–145.

[38] M. Kazemzadeh, M. Martinez-Calderon, W. Xu, L. W. Chamley, C. L. Hisey, and N. G. Broderick, "Cascaded deep convolutional neural networks as improved methods of preprocessing raman spectroscopy data," *Analytical Chemistry*, vol. 94, no. 37, pp. 12 907–12 918, 2022.

[39] M. Bilal, G. Ali, M. W. Iqbal, M. Anwar, M. S. A. Malik, and R. A. Kadir, "Auto-prep: efficient and automated data preprocessing pipeline," *IEEE Access*, vol. 10, pp. 107 764–107 784, 2022.

[40] J. Chen, D. Chen, and G. Liu, "Using temporal convolution network for remaining useful lifetime prediction," *Engineering Reports*, vol. 3, no. 3, p. e12305, 2021.

[41] C. Yang, K. Zhou, and J. Liu, "Supergraph: Spatial-temporal graph-based feature extraction for rotating machinery diagnosis," *IEEE Transactions on Industrial Electronics*, vol. 69, no. 4, pp. 4167–4176, 2021.

[42] G. Luo, H. Zhang, Q. Yuan, J. Li, and F.-Y. Wang, "Estnet: embedded spatial-temporal network for modeling traffic flow dynamics," *IEEE transactions on intelligent trans-*

portation systems, vol. 23, no. 10, pp. 19 201–19 212, 2022.

[43] P. Song and C. Zhao, "Slow down to go better: A survey on slow feature analysis," *IEEE Transactions on Neural Networks and Learning Systems*, 2022.

[44] G. Zhang, L. Tang, Z. Liu, L. Zhou, Y. Liu, Z. Jiang, J. Chen, and S. Sun, "Enhanced features in principal component analysis with spatial and temporal windows for damage identification," *Inverse Problems in Science and Engineering*, vol. 29, no. 13, pp. 2877–2894, 2021.

[45] C. Cortes and V. Vapnik, "Support vector networks," *Machine Learning*, vol. 20, pp. 273–297, 1995.

[46] K. Altun, B. Barshan, and O. Tunçel, "Comparative study on classifying human activities with miniature inertial and magnetic sensors," *Pattern Recognition*, vol. 43, no. 10, pp. 3605–3620, 2010.

[47] R. E. Schapire, "Explaining AdaBoost," in *Empirical Inference*, B. Schölkopf, Z. Luo, and V. Vovk, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2013, pp. 37–52. [Online]. Available: https://link.springer.com/10.1007/978-3-642-41136-6_5

[48] O. Bousquet, U. Von Luxburg, and G. Rätsch, Eds., *Advanced Lectures on Machine Learning*, ser. Lecture Notes in Computer Science. Berlin, Heidelberg: Springer Berlin Heidelberg, 2004, vol. 3176. [Online]. Available: http://link.springer.com/10.1007/b100712

[49] T. K. Ho, "Random decision forests," in *Proceedings of 3rd international conference on document analysis and recognition*, vol. 1. IEEE, 1995, pp. 278–282. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/598994/

[50] L. Breiman and J. H. Friedman, "Predicting multivariate responses in multiple linear regression," *Journal of the Royal Statistical Society Series B: Statistical Methodology*, vol. 59, no. 1, pp. 3–54, 1997.

[51] G. Sarailidis, T. Wagener, and F. Pianosi, "Integrating scientific knowledge into machine learning using interactive decision trees," *Computers & Geosciences*, vol. 170, p. 105248,

2023.

[52] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," *Proceedings of the IEEE*, vol. 86, no. 11, pp. 2278–2324, 1998, publisher: Ieee. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/726791/

[53] L. Eren, T. Ince, and S. Kiranyaz, "A generic intelligent bearing fault diagnosis system using compact adaptive 1d cnn classifier," *Journal of Signal Processing Systems*, vol. 91, pp. 179–189, 2019.

[54] G. Li, C. Deng, J. Wu, X. Xu, X. Shao, and Y. Wang, "Sensor data-driven bearing fault diagnosis based on deep convolutional neural networks and s-transform," *Sensors*, vol. 19, no. 12, p. 2750, 2019.

[55] Z. Tang, Z. Chen, Y. Bao, and H. Li, "Convolutional neural network-based data anomaly detection method using multiple information for structural health monitoring," *Structural Control and Health Monitoring*, vol. 26, no. 1, p. e2296, 2019.

[56] G. Memis and M. Sert, "Multimodal classification of obstructive sleep apnea using feature level fusion," in *2017 IEEE 11th international conference on semantic computing (ICSC)*. IEEE, 2017, pp. 85–88.

[57] M. Ravan, "A machine learning approach using eeg signals to measure sleep quality," *AIMS Electron. Electr. Eng*, vol. 3, no. 4, p. 347, 2019.

[58] S. S. Rajagopalan, S. Bhardwaj, R. Panda, V. R. Reddam, C. Ganne, R. Kenchaiah, R. C. Mundlamuri, T. Kandavel, K. K. Majumdar, S. Parthasarathy *et al.*, "Machine learning detects eeg microstate alterations in patients living with temporal lobe epilepsy," *Seizure*, vol. 61, pp. 8–13, 2018.

[59] A. Sarkar, A. Singh, and R. Chakraborty, "A deep learning-based comparative study to track mental depression from eeg data," *Neuroscience Informatics*, vol. 2, no. 4, p. 100039, 2022.

[60] S. Katsigiannis and N. Ramzan, "Dreamer: A database for emotion recognition through eeg and ecg signals from wireless low-cost off-the-shelf devices," *IEEE Journal of Biomedical and Health Informatics*, vol. 22, no. 1, pp. 98–107, 2017.

[61] M. Pisipati and A. Nandy, "Human emotion recognition using eeg signal in music listening," in *2021 IEEE 18th India Council International Conference (INDICON)*. IEEE, 2021, pp. 1–6.

[62] N. H. Saputra and N. Nafi'Iyah, "Identification of human stress based on eeg signals using machine learning," in *2022 1st International Conference on Information System & Information Technology (ICISIT)*. IEEE, 2022, pp. 176–180.

[63] A. L. Goldberger *et al.*, "Physiobank, physiotoolkit, and physionet: Components of a new research resource for complex physiologic signals," *Circulation*, vol. 101, no. 23, pp. e215–e220, 2000.

[64] W.-L. Zheng and B.-L. Lu, "Investigating critical frequency bands and channels for eeg-based emotion recognition with deep neural networks," *IEEE Transactions on Autonomous Mental Development*, vol. 7, no. 3, pp. 162–175, 2015.

[65] T. Tian, L. Wang, M. Luo, Y. Sun, and X. Liu, "Resnet-50 based technique for eeg image characterization due to varying environmental stimuli," *Computer Methods and Programs in Biomedicine*, vol. 225, p. 107092, 2022.

[66] B. Kemp, A. H. Zwinderman, B. Tuk, H. A. Kamphuisen, and J. J. L. Oberyé, "Analysis of a sleep-dependent neuronal feedback loop: the slow-wave microcontinuity of the eeg," *IEEE Transactions on Biomedical Engineering*, vol. 47, no. 9, pp. 1185–1194, 2000.

[67] C. Nash, R. Nair, and S. M. Naqvi, "Machine learning and adhd mental health detection-a short survey," in *2022 25th International Conference on Information Fusion (FUSION)*. IEEE, 2022, pp. 1–8.

[68] M. K. Bashar, I. Chiaki, and H. Yoshida, "Human identification from brain eeg signals using advanced machine learning method eeg-based biometrics," in *2016 IEEE EMBS*

Conference on Biomedical Engineering and Sciences (IECBES). IEEE, 2016, pp. 475–479.

[69] L. D. Sharma, V. K. Bohat, M. Habib, A.-Z. Ala'M, H. Faris, and I. Aljarah, "Evolutionary inspired approach for mental stress detection using eeg signal," *Expert systems with applications*, vol. 197, p. 116634, 2022.

[70] G. Dornhege, B. Blankertz, G. Curio, and K. Müller, "Boosting bit rates in non-invasive eeg single-trial classifications by feature combination and multi-class paradigms," in *IEEE Transactions on Biomedical Engineering*, 2004, vol. 51, no. 6, pp. 993–1002.

[71] R. Chatterjee, T. Bandyopadhyay, D. K. Sanyal, and D. Guha, "Dimensionality reduction of eeg signal using fuzzy discernibility matrix," in *2017 10th International Conference on Human System Interactions (HSI)*. IEEE, 2017, pp. 131–136.

[72] R. Chatterjee, T. Maitra, S. H. Islam, M. M. Hassan, A. Alamri, and G. Fortino, "A novel machine learning based feature selection for motor imagery eeg signal classification in internet of medical things environment," *Future Generation Computer Systems*, vol. 98, pp. 419–434, 2019.

[73] V. Rakshith, V. Apoorv, N. Akarsh, K. Arjun, B. Krupa, M. Pratima, and A. Vedamurthachar, "A novel approach for the identification of chronic alcohol users from ecg signals," in *TENCON 2017-2017 IEEE Region 10 Conference*. IEEE, 2017, pp. 1321–1326.

[74] R. Supakar, P. Satvaya, and P. Chakrabarti, "A deep learning based model using rnn-lstm for the detection of schizophrenia from eeg data," *Computers in Biology and Medicine*, vol. 151, p. 106225, 2022.

[75] W. Li, L. Shao, W. Wang, H. Li, X. Wang, Y. Li, W. Li, T. Jones, and D. Zhang, "Air quality improvement in response to intensified control strategies in beijing during 2013–2019," *Science of the Total Environment*, vol. 744, p. 140776, 2020.

[76] F. Cui, Y. Yue, Y. Zhang, Z. Zhang, and H. S. Zhou, "Advancing biosensors with machine learning," *ACS sensors*, vol. 5, no. 11, pp. 3346–3364, 2020.

[77] S. Wold, K. Esbensen, and P. Geladi, "Principal component analysis," *Chemometrics and intelligent laboratory systems*, vol. 2, no. 1-3, pp. 37–52, 1987.

[78] T. H. Cormen, C. E. Leiserson, R. L. Rivest, and C. Stein, *Introduction to algorithms*. MIT press, 2022.

[79] S. Hochreiter, "The vanishing gradient problem during learning recurrent neural nets and problem solutions." *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, vol. 6, no. 2, pp. 107–116, 1998. [Online]. Available: http://dblp.uni-trier.de/db/journals/ijufks/ijufks6.html#Hochreiter98

[80] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Computation*, vol. 9, no. 8, pp. 1735–1780, 1997.

[81] K. Sako, B. N. Mpinda, and P. C. Rodrigues, "Neural networks for financial time series forecasting," *Entropy*, vol. 24, no. 5, p. 657, 2022.

[82] G. Dudek, S. Smyl, and P. Pełka, "Recurrent neural networks for forecasting time series with multiple seasonality: A comparative study," *arXiv preprint arXiv:2203.09170*, 2022.

[83] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, Ł. Kaiser, and I. Polosukhin, "Attention is all you need," *Advances in neural information processing systems*, vol. 30, 2017.

[84] H. Zhou, S. Zhang, J. Peng, S. Zhang, J. Li, H. Xiong, and W. Zhang, "Informer: Beyond efficient transformer for long sequence time-series forecasting," in *Proceedings of the AAAI conference on artificial intelligence*, vol. 35, no. 12, 2021, pp. 11 106–11 115.

[85] H. Wu, J. Xu, J. Wang, and M. Long, "Autoformer: Decomposition transformers with auto-correlation for long-term series forecasting," *Advances in neural information processing systems*, vol. 34, pp. 22 419–22 430, 2021.

[86] B. Lim, S. Ö. Arık, N. Loeff, and T. Pfister, "Temporal fusion transformers for interpretable multi-horizon time series forecasting," *International Journal of Forecasting*, vol. 37, no. 4, pp. 1748–1764, 2021.

[87] Z. Lu, B. Ozek, and S. Kamarthi, "Transformer encoder with multiscale deep learning for pain classification using physiological signals," *Frontiers in Physiology*, vol. 14, p. 1294577, 2023.

[88] Z. Wang, W. Wu, and C. Zeng, "Physioformer: Integrating multimodal physiological signals and symbolic regression for explainable affective state prediction," *arXiv preprint arXiv:2410.11376*, 2024.

[89] P. Schmidt, A. Reiss, R. Duerichen, C. Marberger, and K. Van Laerhoven, "Introducing wesad, a multimodal dataset for wearable stress and affect detection," in *Proceedings of the 20th ACM international conference on multimodal interaction*, 2018, pp. 400–408.

[90] L. He and W. Luo, "3d-convlstmnet: A deep spatio-temporal model for traffic flow prediction," in *2022 23rd IEEE International Conference on Mobile Data Management (MDM)*. IEEE, 2022, pp. 147–152.

[91] T. Li, A. Ni, C. Zhang, G. Xiao, and L. Gao, "Short-term traffic congestion prediction with conv–bilstm considering spatio-temporal features," *IET Intelligent Transport Systems*, vol. 14, no. 14, pp. 1978–1986, 2020.

[92] S. Abirami, P. Chitra, R. Madhumitha, and S. R. Kesavan, "Hybrid spatio-temporal deep learning framework for particulate matter (pm 2.5) concentration forecasting," in *2020 International conference on innovative trends in information technology (ICITIIT)*. IEEE, 2020, pp. 1–6.

[93] A. Arbelle, S. Cohen, and T. R. Raviv, "Dual-task convlstm-unet for instance segmentation of weakly annotated microscopy videos," *IEEE Transactions on Medical Imaging*, vol. 41, no. 8, pp. 1948–1960, 2022.

[94] H. Wang, C. Lu, Q. Zhang, Z. Hu, X. Yuan, P. Zhang, and W. Liu, "A novel sleep staging network based on multi-scale dual attention," *Biomedical Signal Processing and Control*, vol. 74, p. 103486, 2022.

[95] D. Alvarez-Estevez and R. M. Rijsman, "Inter-database validation of a deep learning approach for automatic sleep scoring," *PloS one*, vol. 16, no. 8, p. e0256111, 2021.

[96] W. Qu, C.-H. Kao, H. Hong, Z. Chi, R. Grunstein, C. Gordon, and Z. Wang, "Single-channel eeg based insomnia detection with domain adaptation," *Computers in biology and medicine*, vol. 139, p. 104989, 2021.

[97] D. S. B. A. Hamid, S. Goyal, and P. Bedi, "Integration of deep learning for improved diagnosis of depression using eeg and facial features," *Materials Today: Proceedings*, vol. 80, pp. 1965–1969, 2023.

[98] A. Zarei, H. Beheshti, and B. M. Asl, "Detection of sleep apnea using deep neural networks and single-lead ecg signals," *Biomedical Signal Processing and Control*, vol. 71, p. 103125, 2022.

[99] P. Tormene, T. Giorgino, S. Quaglini, and M. Stefanelli, "Matching incomplete time series with dynamic time warping: An algorithm and an application to post-stroke rehabilitation," *Artificial Intelligence in Medicine*, vol. 45, no. 1, pp. 11–34, 2008.

[100] B. Schölkopf, A. Smola, and K.-R. Müller, "Comparing support vector machines with gaussian kernels to radial basis function classifiers," *IEEE Transactions on Signal Processing*, 1997.

[101] G. E. Hinton and R. R. Salakhutdinov, "Reducing the dimensionality of data with neural networks," *Science*, vol. 313, no. 5786, pp. 504–507, 2006.

[102] I. T. Jolliffe, *Principal component analysis*. Springer Series in Statistics, 2002.

[103] M. Cuturi, "Fast global alignment kernels," in *ICML*, 2011.

[104] J. Lines and A. Bagnall, "Time series classification with ensembles of elastic distance measures," *Data Mining and Knowledge Discovery*, vol. 26, no. 2, pp. 565–592, 2012.

[105] D. R. Cox, "The regression analysis of binary sequences," *Journal of the Royal Statistical Society Series B: Statistical Methodology*, vol. 20, no. 2, pp. 215–232, 1958, publisher: Oxford University Press. [Online]. Available: https://academic.oup.com/jrsssb/article-abstract/20/2/215/7027376

[106] M. D. Salawu, M. O. Arowolo, S. O. Abdulsalam, R. M. Isiaka, B. Jimada-Ojuolape, M. L. Olumide, and K. A. Gbolagade, "A chi-square-svm based pedagogical rule extraction method for microarray data analysis," *International Journal of Advances in Applied Sciences*, vol. 9, no. 2, pp. 93–100, 2020.

[107] T. T. Le, W. Fu, and J. H. Moore, "Scaling tree-based automated machine learning to biomedical big data with a feature set selector," *Bioinformatics*, vol. 36, no. 1, pp. 250–256, 2020, publisher: Oxford University Press. [Online]. Available: https://academic.oup.com/bioinformatics/article-abstract/36/1/250/5511404

[108] A. Kumar, K. Sharma, and A. Sharma, "Genetically optimized Fuzzy C-means data clustering of IoMT-based biomarkers for fast affective state recognition in intelligent edge analytics," *Applied Soft Computing*, vol. 109, p. 107525, 2021, publisher: Elsevier. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1568494621004488

[109] Q. Xu, T. L. Nwe, and C. Guan, "Cluster-based analysis for personalized stress evaluation using physiological signals," *IEEE journal of biomedical and health informatics*, vol. 19, no. 1, pp. 275–281, 2014.

[110] M. M. Taye, "Theoretical understanding of convolutional neural network: concepts, architectures, applications, future directions," *Computation*, vol. 11, no. 3, p. 52, 2023.

[111] S. Chakraborty, S. Aich, M.-i. Joo, M. Sain, and H.-C. Kim, "A multichannel convolutional neural network architecture for the detection of the state of mind using physiological signals from wearable devices," *Journal of healthcare engineering*, vol. 2019, 2019, publisher: Hindawi Limited. [Online]. Available: https://downloads.hindawi.com/journals/jhe/2019/5397814.pdf

[112] F. Karim, S. Majumdar, H. Darabi, and S. Harford, "Multivariate lstm-fcns for time series classification," *Neural networks*, vol. 116, pp. 237–245, 2019, publisher: Elsevier. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0893608019301200

[113] K. L. Wolf, S. T. Lam, J. K. McKeen, G. R. Richardson, M. Van Den Bosch, and A. C. Bardekjian, "Urban trees and human health: A scoping review," *International journal of environmental research and public health*, vol. 17, no. 12, p. 4371, 2020.

[114] G. Grilli and S. Sacchelli, "Health benefits derived from forest: A review," *International journal of environmental research and public health*, vol. 17, no. 17, p. 6125, 2020.

[115] J. LaFantasie, "Citizen science and public health-can ebird data inform relationships between public health and access to biodiversity?" *medRxiv*, pp. 2022–01, 2022.

[116] M. J. Maes, M. Pirani, E. R. Booth, C. Shen, B. Milligan, K. E. Jones, and M. B. Toledano, "Benefit of woodland and other natural environments for adolescents' cognition and mental health," *Nature sustainability*, vol. 4, no. 10, pp. 851–858, 2021.

[117] P.-S. Yeon, I.-O. Kim, S.-N. Kang, N.-E. Lee, G.-Y. Kim, G.-M. Min, C.-Y. Chung, J.-S. Lee, J.-G. Kim, and W.-S. Shin, "Effects of urban forest therapy program on depression patients," *International journal of environmental research and public health*, vol. 20, no. 1, p. 507, 2022.

[118] P. Liu, M. Liu, T. Xia, Y. Wang, and P. Guo, "The relationship between landscape metrics and facial expressions in 18 urban forest parks of northern china," *Forests*, vol. 12, no. 12, p. 1619, 2021.

[119] M. Geneshka, C. J. McClean, S. Gilbody, J. Cruz, and P. Coventry, "The future is green. integrating green and blue space data from european urban atlas into uk biobank," *medRxiv*, pp. 2022–05, 2022.

[120] M. P. White, L. R. Elliott, J. Grellier, T. Economou, S. Bell, G. N. Bratman, M. Cirach, M. Gascon, M. L. Lima, M. Lõhmus *et al.*, "Associations between green/blue spaces and mental health across 18 countries," *Scientific reports*, vol. 11, no. 1, p. 8903, 2021.

[121] Z. Li, Z. He, Y. Zhang, S. Jin, X. Wang, J. Zhu, and S. Liu, "Impact of greenspace exposure on residents' mental health: A case study of nanjing city," *Progress in Geography*, vol. 39, no. 5, pp. 779–791, 2020.

[122] B. Qin, W. Zhu, J. Wang, and Y. Peng, "Understanding the relationship between neighbourhood green space and mental wellbeing: A case study of beijing, china," *Cities*, vol. 109, p. 103039, 2021.

[123] B.-X. Huang, S.-C. Chiou, and W.-Y. Li, "Accessibility and street network characteristics of urban public facility spaces: Equity research on parks in fuzhou city based on gis and space syntax model," *Sustainability*, vol. 12, no. 9, p. 3618, 2020.

[124] D. Luo, H. Heacock *et al.*, "Access to green space and median household income in metro vancouver cities," *BCIT Environmental Public Health Journal*, 2020.

[125] Y. Zhang, Q. Wu, L. Wu, and Y. Li, "Measuring community green inequity: A fine-scale assessment of beijing urban area," *Land*, vol. 10, no. 11, p. 1197, 2021.

[126] P. Chaitanya, E. Upadhyay, D. D. Singh, and V. Singh, "Effective contribution of air pollutants to physiological and psychological human diseases: A systematic review." *Nature Environment & Pollution Technology*, vol. 21, no. 4, 2022.

[127] K. R. Cromar, L. A. Gladson, E. A. Hicks, B. Marsh, and G. Ewart, "Excess morbidity and mortality associated with air pollution above american thoracic society recommended standards, 2017–2019," *Annals of the American Thoracic Society*, vol. 19, no. 4, pp. 603–613, 2022.

[128] Y.-C. Yang, J.-W. Li, B. Sun, Y.-L. Chen, S.-Q. Shen, and C. Yuan, "Cardiovascular and cerebrovascular diseases induced by air pollution," *Zhongguo yi xue ke xue Yuan xue bao. Acta Academiae Medicinae Sinicae*, vol. 44, no. 2, pp. 318–323, 2022.

[129] H. Liu, Y. Chen, and L. Ma, "Effect of time-varying exposure to air pollution on subjective well-being," *Journal of Cleaner Production*, vol. 281, p. 125364, 2021.

[130] P. D. Katoto, A. S. Brand, B. Bakan, P. M. Obadia, C. Kuhangana, T. Kayembe-Kitenge, J. P. Kitenge, C. B. L. Nkulu, J. Vanoirbeek, T. S. Nawrot *et al.*, "Acute and chronic exposure to air pollution in relation with incidence, prevalence, severity and mortality of covid-19: a rapid systematic review," *Environmental Health*, vol. 20, pp. 1–21, 2021.

[131] N. Ali and F. Islam, "The effects of air pollution on covid-19 infection and mortality—a review on recent evidence," *Frontiers in public health*, vol. 8, p. 580057, 2020.

[132] A. Khajavi, S. S. T. Zadeh, F. Azizi, R. D. Brook, H. Abdi, F. Zayeri, and F. Hadaegh, "Impact of short-and long-term exposure to air pollution on blood pressure: A two-decade population-based study in tehran," *International Journal of Hygiene and Environmental Health*, vol. 234, p. 113719, 2021.

[133] I. Chandra, K. Nisa, and E. Rosdiana, "Preliminary study: health risk analysis of pm2. 5 and pm10 mass concentrations in bandung metropolitan," in *IOP Conference Series: Earth and Environmental Science*. IOP Publishing, 2021, p. 012049.

[134] S. Baquie and H. Fuje, *Vulnerability to poverty following extreme weather events in Malawi*. The World Bank, 2020.

[135] M. Iannuccilli, G. Bartolini, G. Betti, A. Crisci, D. Grifoni, B. Gozzini, A. Messeri, M. Morabito, C. Tei, T. Torrigiani Malaspina *et al.*, "Extreme precipitation events and their relationships with circulation types in italy," *International Journal of Climatology*, vol. 41, no. 10, pp. 4769–4793, 2021.

[136] A. Sulikowska and A. Wypych, "Seasonal variability of trends in regional hot and warm temperature extremes in europe," *Atmosphere*, vol. 12, no. 5, p. 612, 2021.

[137] M. J. Hewer, "Determining the effect of extreme weather events on human participation in recreation and tourism: a case study of the toronto zoo," *Atmosphere*, vol. 11, no. 1, p. 99, 2020.

[138] K. Bhattacharjee, M. Chen, and A. Dasgupta, "Privacy-preserving data visualization: reflections on the state of the art and research opportunities," in *Computer Graphics Forum*. Wiley Online Library, 2020, pp. 675–692.

[139] J. Martinsson, E. L. Zec, D. Gillblad, and O. Mogren, "Adversarial representation learning for synthetic replacement of private attributes," in *2021 IEEE International Conference on Big Data (Big Data)*. IEEE, 2021, pp. 1291–1299.

[140] Y. Tachioka, "Privacy preservation satisfying utility requirements based on multi-objective optimization," in *2022 Joint 12th International Conference on Soft Computing and Intelligent Systems and 23rd International Symposium on Advanced Intelligent Systems (SCIS&ISIS)*. IEEE, 2022, pp. 1–4.

[141] A. Yala, V. Quach, H. Esfahanizadeh, R. G. D'Oliveira, K. R. Duffy, M. Médard, T. S. Jaakkola, and R. Barzilay, "Syfer: Neural obfuscation for private data release," *arXiv preprint arXiv:2201.12406*, 2022.

[142] Y.-F. Ge, W.-J. Yu, J. Cao, H. Wang, Z.-H. Zhan, Y. Zhang, and J. Zhang, "Distributed memetic algorithm for outsourced database fragmentation," *IEEE Transactions on Cybernetics*, vol. 51, no. 10, pp. 4808–4821, 2020.

[143] M. Boedihardjo, T. Strohmer, and R. Vershynin, "Covariance's loss is privacy's gain: Computationally efficient, private and accurate synthetic data," *Foundations of Computational Mathematics*, pp. 1–48, 2022.

[144] B. van Deenen, P. Nauts, R. Trietsch, and B. Voorn, "Towards privacy by design for data with strm privacy," *Data Engineering*, p. 62, 2022.

[145] F. Sarzotti, "Self-monitoring of emotions and mood using a tangible approach," *Computers*, vol. 7, no. 1, p. 7, 2018.

[146] V. Roto, J. Häkkilä, K. Väänänen-Vainio-Mattila, O. Juhlin, T. Olsson, and E. Hvannberg, *Proceedings of the 8th Nordic Conference on Human-Computer Interaction: Fun, Fast, Foundational*. ACM, 2014.

[147] F.-T. Sun, C. Kuo, H.-T. Cheng, S. Buthpitiya, P. Collins, and M. Griss, "Activity-aware mental stress detection using physiological sensors," in *Mobile Computing, Applications, and Services: Second International ICST Conference, MobiCASE 2010, Santa Clara, CA, USA, October 25-28, 2010, Revised Selected Papers 2*. Springer, 2012, pp. 282–301.

[148] A. Mishra and R. Yu, "Urban planning & mental health: How can we design healthier cities for our aging population to improve mental health outcomes?" *University of*

Western Ontario Medical Journal, vol. 88, no. 2, pp. 33–36, 2020.

[149] D. R. Sahoo, T. Neate, Y. Tokuda, J. Pearson, S. Robinson, S. Subramanian, and M. Jones, "Tangible drops: a visio-tactile display using actuated liquid-metal droplets," in *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, 2018, pp. 1–14.

[150] T.-D. Tran, J. Kim, N.-H. Ho, H.-J. Yang, S. Pant, S.-H. Kim, and G.-S. Lee, "Stress analysis with dimensions of valence and arousal in the wild," *Applied Sciences*, vol. 11, no. 11, p. 5194, 2021.

[151] G. Giannakakis, D. Grigoriadis, K. Giannakaki, O. Simantiraki, A. Roniotis, and M. Tsiknakis, "Review on psychological stress detection using biosignals," *IEEE Transactions on Affective Computing*, vol. 13, no. 1, pp. 440–460, 2019, publisher: IEEE. [Online]. Available: https://ieeexplore.ieee.org/iel7/5165369/5520654/08758154.pdf

[152] K. Tzevelekakis, Z. Stefanidi, and G. Margetis, "Real-time stress level feedback from raw ecg signals for personalised, context-aware applications using lightweight convolutional neural network architectures," *Sensors*, vol. 21, no. 23, p. 7802, 2021, publisher: MDPI. [Online]. Available: https://www.mdpi.com/1424-8220/21/23/7802

[153] E. Dogan, C. Sander, X. Wagner, U. Hegerl, and E. Kohls, "Smartphone-based monitoring of objective and subjective data in affective disorders: where are we and where are we going? Systematic review," *Journal of medical Internet research*, vol. 19, no. 7, p. e262, 2017, publisher: JMIR Publications Toronto, Canada. [Online]. Available: https://www.jmir.org/2017/7/e262/

[154] P. Jakobsen, E. Garcia-Ceja, M. Riegler, L. A. Stabell, T. Nordgreen, J. Torresen, O. B. Fasmer, and K. J. Oedegaard, "Applying machine learning in motor activity time series of depressed bipolar and unipolar patients," *bioRxiv*, pp. 2020–04, 2020, publisher: Cold Spring Harbor Laboratory. [Online]. Available: https://www.biorxiv.org/content/10.1101/2020.04.07.029561.abstract

[155] A. Ortiz and P. Grof, "Electronic monitoring of self-reported mood: the return of the

subjective?" *International Journal of Bipolar Disorders*, vol. 4, no. 1, p. 28, Dec. 2016. [Online]. Available: http://journalbipolardisorders.springeropen.com/articles/10.1186/s40345-016-0069-x

[156] E. Garcia-Ceja, M. Riegler, P. Jakobsen, J. Tørresen, T. Nordgreen, K. J. Oedegaard, and O. B. Fasmer, "Depresjon: a motor activity database of depression episodes in unipolar and bipolar patients," in *Proceedings of the 9th ACM Multimedia Systems Conference*. Amsterdam Netherlands: ACM, Jun. 2018, pp. 472–477. [Online]. Available: https://dl.acm.org/doi/10.1145/3204949.3208125

[157] B. Szakonyi, I. Vassányi, E. Schumacher, and I. Kósa, "Efficient methods for acute stress detection using heart rate variability data from Ambient Assisted Living sensors," *BioMedical Engineering OnLine*, vol. 20, no. 1, p. 73, Dec. 2021. [Online]. Available: https://biomedical-engineering-online.biomedcentral.com/articles/10.1186/s12938-021-00911-6

[158] E. Garcia-Ceja, M. Riegler, P. Jakobsen, J. Tørresen, T. Nordgreen, K. J. Oedegaard, and O. B. Fasmer, "Depresjon: a motor activity database of depression episodes in unipolar and bipolar patients," in *Proceedings of the 9th ACM Multimedia Systems Conference*. Amsterdam Netherlands: ACM, Jun. 2018, pp. 472–477. [Online]. Available: https://dl.acm.org/doi/10.1145/3204949.3208125

[159] C. E. Galvan-Tejada, L. A. Zanella-Calzada, H. Gamboa-Rosales, J. I. Galván-Tejada, N. M. Chávez-Lamas, M. Gracia-Cortés, R. Magallanes-Quintanar, and J. M. Celaya-Padilla, "Depression episodes detection in unipolar and bipolar patients: a methodology with feature extraction and feature selection with genetic algorithms using activity motion signal as information source," *Mobile Information Systems*, vol. 2019, 2019, publisher: Hindawi. [Online]. Available: https://www.hindawi.com/journals/misy/2019/8269695/abs/

[160] J. G. Rodríguez-Ruiz, C. E. Galván-Tejada, L. A. Zanella-Calzada, J. M. Celaya-Padilla, J. I. Galván-Tejada, H. Gamboa-Rosales, H. Luna-García, R. Magallanes-Quintanar, and M. A. Soto-Murillo, "Comparison of night, day and 24 h motor activity data for

the classification of depressive episodes," *Diagnostics*, vol. 10, no. 3, p. 162, 2020, publisher: MDPI. [Online]. Available: https://www.mdpi.com/2075-4418/10/3/162

[161] L. B. Leng, L. B. Giin, and W.-Y. Chung, "Wearable driver drowsiness detection system based on biomedical and motion sensors," in *2015 IEEE SENSORS*. IEEE, 2015, pp. 1–4. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/7370355/

[162] P. Jakobsen, E. Garcia-Ceja, M. Riegler, L. A. Stabell, T. Nordgreen, J. Torresen, O. B. Fasmer, and K. J. Oedegaard, "Applying machine learning in motor activity time series of depressed bipolar and unipolar patients," *bioRxiv*, pp. 2020–04, 2020, publisher: Cold Spring Harbor Laboratory. [Online]. Available: https://www.biorxiv.org/content/10.1101/2020.04.07.029561.abstract

[163] E. Garcia-Ceja, M. Riegler, P. Jakobsen, J. Tørresen, T. Nordgreen, K. J. Oedegaard, and O. B. Fasmer, "Depresjon: a motor activity database of depression episodes in unipolar and bipolar patients," in *Proceedings of the 9th ACM Multimedia Systems Conference*. Amsterdam Netherlands: ACM, Jun. 2018, pp. 472–477. [Online]. Available: https://dl.acm.org/doi/10.1145/3204949.3208125

[164] C. Cortes and V. Vapnik, "Support-vector networks," *Machine Learning*, vol. 20, no. 3, pp. 273–297, Sep. 1995. [Online]. Available: http://link.springer.com/10.1007/BF00994018

[165] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg *et al.*, "Scikit-learn: Machine learning in python," *the Journal of machine Learning research*, vol. 12, pp. 2825–2830, 2011.

[166] B. Chaix, "Mobile sensing in environmental health and neighborhood research," *Annual review of public health*, vol. 39, pp. 367–384, 2018.

[167] M. J. Nieuwenhuijsen, D. Donaire-Gonzalez, M. Foraster, D. Martinez, and A. Cisneros, "Using personal sensors to assess the exposome and acute health effects," *International Journal of Environmental Research and Public Health*, vol. 11, no. 8, pp. 7805–7819, 2014.

[168] T. Johnson, E. Kanjo, and K. Woodward, "Digitalexposome: quantifying impact of urban environment on wellbeing using sensor fusion and deep learning," *Computational Urban Science*, vol. 3, no. 1, p. 14, 2023.

[169] T. Johnson and E. Kanjo, "Urban wellbeing: A portable sensing approach to unravel the link between environment and mental wellbeing," *IEEE Sensors Letters*, vol. 7, no. 3, pp. 1–4, 2023.

[170] ——, "Sensor fusion and the city: visualisation and aggregation of environmental & wellbeing data," in *2021 IEEE International Smart Cities Conference (ISC2)*. IEEE, 2021, pp. 1–4.

[171] M. Basner, W. Babisch, A. Davis, M. Brink, C. Clark, S. Janssen, and S. Stansfeld, "Auditory and non-auditory effects of noise on health," *The lancet*, vol. 383, no. 9925, pp. 1325–1332, 2014.

[172] R. Dai, C. Lu, L. Yun, E. Lenze, M. Avidan, and T. Kannampallil, "Comparing stress prediction models using smartwatch physiological signals and participant self-reports," *Computer Methods and Programs in Biomedicine*, vol. 208, p. 106207, 2021.

[173] T. Johnson, E. Kanjo, and K. Woodward, "Digitalexposome: Quantifying the urban environment influence on wellbeing based on real-time multi-sensor fusion and deep belief network," in *Proceedings of the 2019 IEEE International Conference on Pervasive Computing and Communications Workshops (PerCom Workshops)*. IEEE, 2019, pp. 589–594.

[174] M. Schutte, A. Mark, E. Wenning, and B. Griefahn, "The development of the noise sensitivity questionnaire," *Noise Health*, vol. 9, pp. 15–24, 2007.

[175] A. Dzhambov and D. Dimitrova, "Validating a short bulgarian version of a psychometric instrument for multidimensional noise sensitivity assessment," *Folia Med (Plovdiv)*, vol. 52, pp. 116–125, 2014. [Online]. Available: 116-25.doi:10.2478/folmed-2014-0017

[176] S. Cohen, T. Kamarck, and R. Mermelstein, "A global measure of perceived stress," *Journal of health and social behavior*, pp. 385–396, 1983.

Appendix A

Appendix 01

A.1 Participants Stress Level

ID	Total Score	Stress Level
1	24	moderate stress
2	20	moderate stress
3	10	low stress
4	15	moderate stress
5	23	moderate stress
6	23	moderate stress
7	9	low stress
8	18	moderate stress
9	19	moderate stress
10	14	moderate stress
11	18	moderate stress
12	16	moderate stress
13	21	moderate stress
14	23	moderate stress
15	10	low stress

16	15	moderate stress
17	18	moderate stress
18	24	moderate stress
19	23	moderate stress
20	12	low stress
21	19	moderate stress
22	17	moderate stress
23	16	moderate stress
24	21	moderate stress
25	18	moderate stress
26	28	high stress
27	26	moderate stress
28	22	moderate stress
29	22	moderate stress
30	23	moderate stress
31	15	moderate stress
32	26	moderate stress
33	10	moderate stress
34	22	moderate stress
35	22	moderate stress
36	18	moderate stress
37	20	moderate stress
38	14	moderate stress
39	26	moderate stress
40	26	moderate stress
41	23	moderate stress
42	22	moderate stress
43	15	moderate stress
44	15	moderate stress

45	18	moderate stress
46	26	moderate stress
47	12	low stress
48	17	moderate stress
49	17	moderate stress
50	16	moderate stress
51	16	moderate stress
52	13	low stress
53	21	moderate stress

Table A.1: Participants stress level from the questionnaire.

A.2 Participants Noise Sensitivity Level

ID	Total Score	Noise Sensitivity
1	24	moderate Noise Sensitivity
2	20	moderate Noise Sensitivity
3	10	moderate Noise Sensitivity
4	15	moderate Noise Sensitivity
5	23	moderate Noise Sensitivity
6	23	high Noise Sensitivity
7	9	moderate Noise Sensitivity
8	18	moderate Noise Sensitivity
9	19	moderate Noise Sensitivity
10	14	moderate Noise Sensitivity
11	18	moderate Noise Sensitivity
12	16	moderate Noise Sensitivity
13	21	moderate Noise Sensitivity
14	23	high Noise Sensitivity
15	10	moderate Noise Sensitivity
16	15	moderate Noise Sensitivity
17	18	moderate Noise Sensitivity
18	24	moderate Noise Sensitivity
19	23	high Noise Sensitivity
20	12	moderate Noise Sensitivity
21	19	moderate Noise Sensitivity
22	17	moderate Noise Sensitivity
23	16	moderate Noise Sensitivity
24	21	moderate Noise Sensitivity
25	18	high Noise Sensitivity
26	28	moderate Noise Sensitivity

27	26	moderate Noise Sensitivity
28	22	high Noise Sensitivity
29	22	moderate Noise Sensitivity
30	23	moderate Noise Sensitivity
31	15	moderate Noise Sensitivity
32	26	moderate Noise Sensitivity
33	10	moderate Noise Sensitivity
34	22	moderate Noise Sensitivity
35	22	high Noise Sensitivity
36	18	high Noise Sensitivity
37	20	moderate Noise Sensitivity
38	14	moderate Noise Sensitivity
39	26	high Noise Sensitivity
40	26	moderate Noise Sensitivity
41	23	moderate Noise Sensitivity
42	22	moderate Noise Sensitivity
43	15	moderate Noise Sensitivity
44	15	moderate Noise Sensitivity
45	18	high Noise Sensitivity
46	26	high Noise Sensitivity
47	12	moderate Noise Sensitivity
48	17	moderate Noise Sensitivity
49	17	moderate Noise Sensitivity
50	16	moderate Noise Sensitivity
51	16	moderate Noise Sensitivity
52	13	moderate Noise Sensitivity
53	21	moderate Noise Sensitivity

Table A.2: Participants Noise Sensitivity level from the questionnaire

Appendix B

Appendix 02

B.1 Demographic Questionnaire

Form A.1: Demographic survey questionnaire

Age: What is your age?

- Under 12 years old
- 12-17 years old
- 18-24 years old
- 25-34 years old
- 35-44 years old
- 45-54 years old
- 55-64 years old
- 65-74 years old
- 75 years or older

Ethnic origin: Please specify your ethnicity.

- White
- Hispanic or Latino
- Black or African American
- Native American or American Indian
- Asian / Pacific Islander
- Other

Education: What is the highest degree or level of school you have completed? If currently enrolled, highest degree received.

- No schooling completed
- Nursery school to 8th grade
- Some high school, no diploma
- High school graduate, diploma or the equivalent (for example: GED)
- Some college credit, no degree
- Trade/technical/vocational training
- Associate degree
- Bachelor's degree
- Master's degree
- Professional degree
- Doctorate degree

Marital Status: What is your marital status?

- Single, never married
- Married or domestic partnership

- Widowed
- Divorced
- Separated

Employment Status: Are you currently. . . ?

- Employed for wages
- Self-employed
- Out of work and looking for work
- Out of work but not currently looking for work
- A homemaker
- A student
- Military
- Retired
- Unable to work

B.2 NoiSeQSF Questionnaire

In the following section, we ask you to make statements concerning different noises. Please go through the statements and do not leave out a statement. Please try to put yourself in the respective situation and answer spontaneously without much thinking about it. Please, mark as you see fit the rating category that applies best. Please tick only one rating category per statement. We are interested in your individual view concerning the statements; hence, there are no correct or incorrect responses. If you feel uncertain which of the rating categories applies best to you, mark the category, which best fits your personal view.

Name: ..... Date: ..... Age: .....

Gender (Circle): M F Other: .....

1 = Disagree 2 = Strongly disagree 3 = Agree 4 = Strongly agree

1. I find it hard to relax in a noisy environment	1 2 3 4
2. I need peace and quiet to do difficult work	1 2 3 4
3. For a quiet place to live I would accept other disadvantages	1 2 3 4
4. I find it hard to communicate while it is noisy	1 2 3 4
5. When I am absorbed in a conversation I do not notice whether it is noisy around me	1 2 3 4
6. I can fall asleep even when it is noisy	1 2 3 4
7.Even the slightest noise can prevent me from falling asleep	1 2 3 4
8.It would not bother me to live in a noisy street	1 2 3 4
9. At weekends I prefer quiet surroundings	1 2 3 4
10. I wake up at the slightest noise	1 2 3 4
11. I avoid leisure activities which are loud	1 2 3 4
12. I don't like noisy activities in my residential area	1 2 3 4
13. High noise levels make it hard for me to concentrate on my conversation	1 2 3 4
14. I have no problems to do routine work in a noisy environment	1 2 3 4
15. I need quiet surroundings to be able to work on new tasks	1 2 3 4
TOTAL SCORE	.....

B.3 PERCEIVED STRESS SCALE (PSS)

The questions in this scale ask you about your feelings and thoughts during the last month. In each case, you will be asked to indicate by circling how often you felt or thought a certain way.

Name: ..... Date: ..... Age: .....

Gender (Circle): M F Other: .....

0 = Never 1 = Almost Never 2 = Sometimes 3 = Fairly Often 4 = Very Often

1. In the last month, how often have you been upset because of something that happened unexpectedly?	0 1 2 3 4
2. In the last month, how often have you felt that you were unable to control the important things in your life?	0 1 2 3 4
3. In the last month, how often have you felt nervous and "stressed"?	0 1 2 3 4
4. In the last month, how often have you felt confident about your ability to handle your personal problems?	0 1 2 3 4
5. In the last month, how often have you felt that things were going your way?	0 1 2 3 4
6. In the last month, how often have you found that you could not cope with all the things that you had 4	
7. In the last month, how often have you been able to control irritations in your life?	0 1 2 3 4
8.In the last month, how often have you felt that you were on top of things?	0 1 2 3 4
9. In the last month, how often have you been angered because of things that were outside of your control?	0 1 2 3 4
10. In the last month, how often have you felt difficulties were piling up so high that you could not overcome them?	0 1 2 3 4
TOTAL SCORE	.....

B.4 Ethic Approval

NOTTINGHAM TRENT UNIVERSITY

School of Science and Technology

Non-Invasive Human Ethics Committee

Notification of Decision– 21/22-53

Date: 10/02/2022

Student's Name	Faiza Guerrache
Supervisor's Name	David Joseph Brown
NTU ID	N0045291
Course	PhD
Title	Spatial and Temporal Environment pollution Impact Analysis on People's Wellbeing
Start Date	Upon ethical approval
End Date	20 May 2023

Approved with Conditions – see points below. Before commencing your research, you must incorporate these conditions and resubmit all the application, **highlighting the changes in yellow**, via SST.ethics@ntu.ac.uk for final approval.

Points the applicant needs to address:

Independent Reviewer 1: Approved with Recommendations

- In the Participant statement replace the following in section 2 "but please not the backpack is not heavy" with "the backpack is not heavy"
- Your restrictions on the participant regarding alcohol and exercise should also be in the consent form (if you are sure that these are needed for your research).

Independent Reviewer 2: Approved with Conditions

Specific Comments:

- Section 2.1 and 2.2: What will the participants be expected to do as part of the project?
- Section 2.4, 2.5. and 2.7: Please check your ethics form to ensure that you have answered and ticked/circled 'yes' or 'no' to some of the questions
- Section 4.4: What are you going to be doing with the participants in these rooms? Do they need to be briefing and debriefed?
- Section 4.6: What safety guidelines do participants need to follow?
- Section 6.3: Are you going to store data on a password protected account? If so, please state
- Section 8.5 and 8.6: What permissions are required to use the smartphone application? Are you storing participant data using this platform? Is this internal/external of the university? If so, how are you going to manage sensitive information of participants?

You must discuss any untoward incident, during this project, which results in the completion of an accident claim form in the first instance to your supervisor as a matter of urgency and then via SST.ethics@ntu.ac.uk.

If you have any queries please do not hesitate to contact your project supervisor or alternatively e-mail SST.ethics@ntu.ac.uk.

Version 1 – 08.09.2016