Optimizing Stress Prediction with Wearable Sensor Data: A Comprehensive Approach Through Advanced Machine Learning Techniques

By

Anay Sinhal

Under The Supervision of

Dr. Subhasis Bhattacharjee

Associate Professor, Deptt. of Computer Sci. & Engg.,

Indian Institute of Technology,

Jammu (J&K)

January 2024

DECLARATION

I hereby declare that the research work titled "**Optimizing Stress Prediction with Wearable Sensor Data: A Comprehensive Approach Through Advanced Machine Learning Techniques**" is the result of my own research conducted under the supervision of **Dr. Subhasis Bhattacharjee** at Indian Institute of Technology, Jammu.

This work, completed with over 45 days of dedicated research and engagement, has not been submitted for any degree or examination in this or any other institution.

I affirm that all sources of material used for the thesis have been duly cited, ensuring the originality and integrity of the research.

Anay Sinhal

ACKNOWLEDGEMENT

I am profoundly grateful for the guidance, support, and expertise bestowed upon me by Dr. Subhasis Bhattacharjee, my esteemed project supervisor, whose insights and encouragement were pivotal to the trajectory and success of my research on Stress Prediction using Wearable Sensors. His mentorship was invaluable, and I extend my heartfelt thanks.

My sincere appreciation also goes to Dr. Arpan Gupta, my internal mentor, whose expertise and constructive feedback significantly contributed to the refinement of this project. His support has been instrumental in navigating the challenges encountered during this research.

I am indebted to my college's administrative and academic leadership, including Prof. Dheeraj Sanghi, the Vice Chancellor; Prof. Renu Jain, the Dean; and our Department Coordinators, Prof. Amit Sinhal and Prof. Taruna Sunil. Their unwavering support and the provision of an enriching academic environment have been crucial to my research endeavours.

Special thanks are extended to my peers and colleagues, whose collaborative spirit, stimulating discussions, and shared passion for research enriched this project in countless ways. Their camaraderie and intellectual generosity have been a source of inspiration and motivation throughout this journey.

The resources and facilities provided by JK Lakshmipat University were indispensable to the execution and completion of this study. For this support, I am incredibly grateful.

Finally, I must acknowledge my family and friends' enduring patience, love, and encouragement. Their belief in my capabilities and unwavering support gave me the strength and perseverance required to pursue and achieve my academic goals.

This project stands as a testament to the collective effort, encouragement, and guidance of everyone mentioned above, and to all of them, I offer my deepest gratitude. Their contributions have facilitated the successful completion of this research and profoundly impacted my personal and professional growth.

Anay Sinhal

TABLE OF CONTENTS

ABSTRACT

In the exigent realm of healthcare, the mental well-being of frontline professionals, especially nurses, is paramount. This study pioneers in presenting an exhaustive dataset, meticulously gathered through wearable sensors amid the tumultuous COVID-19 pandemic, to monitor real-time stress. It encapsulates physiological data—electrodermal activity, heart rate, and skin temperature—paired with subjective stress evaluations, offering a holistic perspective on stress dynamics within hospital environments.

Anchored by a methodical literature review that navigates through the landscape of stress detection methodologies, focusing on wearable and IoT technologies, this research advances by dissecting the collected data with sophisticated machine learning algorithms. The dataset, encompassing over 11.5 million entries from 15 nurses across diverse shifts, is a testament to the intricate interplay between physiological responses and occupational stressors.

The analytical journey begins with an elaborate data preprocessing stage, leveraging the Synthetic Minority Over-sampling Technique (SMOTE) to mitigate data imbalances, enriching the dataset for optimal model training. Employing cutting-edge models such as Random Forest, XGBoost, and MLPClassifier culminates in a robust Stacking Classifier that synergizes their predictive prowess, showcasing exceptional accuracy and offering nuanced insights into stress predictors.

The evaluative lens of this research not only celebrates the high accuracy of the models but also delves into the critical analysis of confusion matrices and classification reports, unravelling the significance of the importance of permutation features. This meticulous approach unearths pivotal physiological indicators crucial for stress prediction, steering future explorations towards the most impactful data streams for monitoring stress effectively.

This groundbreaking study not only underscores the transformative potential of wearable technologies in monitoring stress in real time but also highlights the instrumental role of machine learning in decoding the complex narratives of physiological data. By presenting a comprehensive dataset, methodological rigour, and insightful analyses, this research furnishes a blueprint for future endeavours aimed at fortifying healthcare workers' mental health, thereby enhancing their ability to deliver exemplary care.

In essence, this research not only contributes a novel dataset and analytical framework to the scientific community but also paves the way for innovative stress management interventions, ensuring the resilience and well-being of healthcare professionals in their noble mission of care.

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1: INTRODUCTION

1.1 Overview

In the vanguard of healthcare technology, the advent of wearable sensors marks a pivotal shift towards the intimate monitoring of physiological states, a leap of particular significance in stress detection for healthcare professionals. This research taps into this transformative potential, presenting a comprehensive dataset focused on hospital nurses, a demographic whose stress levels have long been of clinical interest yet seldom captured with such granularity and scale. The dataset, emerging amidst the turmoil of the COVID-19 pandemic, offers an unfiltered lens into the physiological signatures of stress captured through the continuous monitoring of vital parameters.

The relevance of this research is manifold, anchored in the stark reality of the healthcare sector's stress epidemic. Nurses, the backbone of clinical operations, are routinely subjected to high-stress situations that directly impact their physical and mental well-being, patient care, and broader healthcare outcomes. By leveraging wearable sensor technology, this study paints a detailed portrait of stress as a dynamic, multifaceted phenomenon, offering insights into its transient spikes and long-term trends across different work shifts and clinical environments.

At the core of this dataset's value is its context—garnered during one of modern history's most challenging periods for healthcare workers. Hence, the data does not merely enumerate stress indicators; it contextualizes them within the high stakes setting of hospitals grappling with a global health crisis. The physiological measures—heart rate, electrodermal activity, and temperature—are not just abstract data points; they are the silent yet eloquent narrators of the nurses' daily experiences, their confrontations with critical incidents, and their moments of resilience or vulnerability.

This study stands on the shoulders of prior scholarly work, integrating and building upon the foundational research in stress monitoring and wearable technologies. It delves into the literature that examines stress from both psychological and physiological perspectives, understanding that stress is not merely a subjective experience but one that leaves tangible traces on the body's biological canvas. This integration of subjective and objective data stands as a testimony to the complexity of stress—a deeply personal yet universally relevant variable.

Moreover, the methodology adopted in this research is a testament to the power of machine learning as an analytical tool capable of dissecting such complexity. Through the application of advanced algorithms, the study not only identifies patterns within the noise but also discerns the nuanced interplay of stress indicators. The machine learning models deployed—Random Forest, XGBoost, and MLPClassifier—were selected for their robustness and their proven track record in predictive analytics. They serve as the analytical engines that power through the dataset, extracting meaningful correlations, predicting stress levels, and offering a predictive gaze into the unseen.

The study's iterative approach to model evaluation and refinement encapsulates the empirical spirit of scientific inquiry. The Stacking Classifier represents a fusion of individual model strengths, showcasing the synergistic potential of machine learning when applied with precision and strategic forethought. The ensemble model not only amplifies accuracy but also underscores the importance of a multifaceted approach in addressing multifaceted problems like stress detection.

Embodied within the dataset and its analysis is the implicit advocacy for a data-informed understanding of workplace stress—a call to action for healthcare systems to integrate such analytics into their operational and support frameworks. The predictive models borne of this research are not academic exercises; they are potential instruments of change, tools that can guide interventions, inform policy, and, ultimately, fortify the well-being of those who stand on the front lines of patient care.

In sum, this overview encapsulates the essence of a research endeavor poised at the intersection of technology, healthcare, and data science. It is a narrative about data's power to illuminate the unseen, about machine learning's capacity to interpret complexity, and, above all, about the imperative to understand and alleviate the stress that pervades the lives of those who dedicate themselves to the service of health.

Figure 1.1: Evolution of Wearable Sensor Technology in Healthcare

1.2 Stress

Stress, inherently a physiological response to perceived challenges, is ubiquitous, manifesting across various facets of life. The body's intrinsic alarm system ignites the fight-or-flight response to immediate threats, thereby orchestrating a cascade of hormonal fluxes and physiological adaptations. While critical for survival, this primal mechanism in chronic manifestation becomes a harbinger of detrimental health outcomes. The literature is replete with studies delineating the multi-dimensional nature of stress, ranging from acute, short-term reactions to persistent, long-term exposures that can precipitate a plethora of adverse health conditions, encompassing cardiovascular disease, diabetes, and a spectrum of mental health disorders.

Research delineates that stress, despite its universal occurrence, varies significantly in its triggers and thresholds, influenced by a confluence of genetic predisposition, psychological resilience, and environmental factors. Therefore, the subjective perception of stress is a conundrum, as its psychological burden is often discordant with biological underpinnings. This discordance poses a challenge in the objective measurement of stress, making it a construct that is as elusive as it is pervasive. However, the quest to quantify and manage stress has galvanized a significant body of research seeking to unveil the biomarkers that can reliably indicate stress levels.

The scholarly discourse extends to the role of cortisol and adrenaline, hormones that serve as biological signifiers of stress, their levels surging in the bloodstream during stressful events. Accompanying physiological responses include elevated heart rate, increased blood pressure, and a heightened state of sensory alertness. While acutely beneficial, these responses can lead to systemic wear and tear when chronically activated, a condition termed allostatic load. It is this chronic activation, devoid of sufficient recovery, that the studies warn against, citing its potential to erode health insidiously.

In summary, the scholarly narrative on stress underscores a critical balance — stress as a catalyst for growth and adaptation against its potential to undermine health when left unchecked. Within this balance, the research seeks to harness the constructive aspects of stress while mitigating its destructive potential.

1.3 Stress in Healthcare

The inherently high-stakes and unpredictable healthcare environment presents a fertile ground for stress. Nurses, in particular, find themselves at the nexus of this stress, navigating the complexities of patient care, administrative pressures, and, often, the emotional toll of human suffering. The literature paints a vivid picture of the healthcare milieu as a crucible of stress, where care demands can stretch the resilience of even the most dedicated professionals.

Research underscores that stress in healthcare is not merely a personal challenge; it has profound implications for the quality of patient care. Nurses under chronic stress are shown to experience cognitive burdens that can impair judgment, reduce empathy, and ultimately compromise patient safety. These findings have catalyzed a growing recognition of the need to address occupational stress within healthcare settings — not as a secondary concern but as a primary determinant of healthcare quality.

The studies converge on the concept that the well-being of healthcare professionals is inextricably linked to the well-being of their patients. They call for an ecosystem of support that extends beyond the individual to encompass the institutional structures that govern healthcare practice. This includes advocating for manageable workloads, fostering a culture of support and resilience, and providing tools and resources that can help nurses cope with the day-to-day stressors of their profession.

Moreover, the research highlights the unique stressors healthcare professionals face during crises, such as the COVID-19 pandemic. The pandemic has amplified the stress levels in healthcare settings exponentially, placing unprecedented demands on nurses and exposing them to high levels of psychological and physical strain. As the literature suggests, the response to this crisis needs to be multifaceted, addressing the immediate need for support and the long-term strategies for building resilience within the healthcare workforce.

In essence, the dialogue within the research community calls for a re-envisioning of healthcare as a field that not only cares for the health of its patients but equally prioritizes the health of those who deliver care. It is a call for systemic change, for interventions that are as much about organizational transformation as they are about individual support, ensuring that the caregivers' stress does not become the overlooked casualty of their noble profession.

1.4 Stress Detection

The recognition and quantification of stress have become increasingly precise with the advent of biometric technology. The recent literature emphasizes the pivotal role of physiological biomarkers in identifying stress levels. The detection hinges on the autonomic nervous system's responses, including variations in heart rate, electrodermal activity, and skin temperature—indicators sensitive to psychological stressors. Innovations in wearable sensor technology have transformed these physiological signals into quantifiable data, enabling continuous, real-time monitoring outside of clinical settings.

The detection of stress is a complex challenge due to the subjective nature of stress experiences and the individual variability in physiological responses. However, machine learning techniques have shown significant promise in deciphering the intricate patterns of these biomarkers, offering a more objective lens through which stress can be measured. By applying algorithms that learn from data, researchers have made strides in developing models that can not only detect the presence of stress but also predict its onset, potentially offering a pre-emptive tool for stress management.

The effectiveness of stress detection models relies heavily on the fidelity of the data captured. Recent studies highlight the importance of high-quality, granular data that wearable sensors can provide. This nuanced understanding of physiological responses has opened the door to personalized stress management interventions, tailoring strategies to the individual's specific stress signatures. Through the continuous evolution of sensor technology and machine learning models, stress detection has progressed from a reactive to a proactive discipline, promising a future where stress can be managed with precision and foresight.

1.5 Stress Detection Techniques

Stress detection techniques have evolved significantly, benefiting from interdisciplinary research in biometrics, psychological assessment, and computational modelling. While helpful in capturing subjective experiences, the traditional method of self-reported questionnaires presents limitations in real-time applicability and objectivity. Modern approaches focus on the direct measurement of physiological signals that reflect the body's response to stressors.

For example, heart rate variability (HRV) has been widely studied for its correlation with stress and emotional states. Electrodermal activity (EDA), which measures changes in skin conductance, offers insights into the sympathetic nervous system's activation, a direct marker of stress. While less commonly used, skin temperature variation provides additional context when analyzing stress responses.

Integrating these physiological measures, advanced algorithms in machine learning have taken centre stage in stress detection. Techniques such as deep learning, which can uncover complex patterns in large datasets, are up and coming. Ensemble methods that combine multiple machine learning models can provide a more robust and accurate prediction by leveraging the strengths of various approaches. With the continuous advancement in sensor accuracy and algorithmic sophistication, these techniques are setting new benchmarks in the field, offering more reliable and actionable insights into stress detection and management.

The data-driven nature of these methods, supported by empirical evidence and the capacity to analyze vast amounts of information, positions them at the forefront of modern stress detection techniques. As the field progresses, these approaches are expected to become more integrated into healthcare practice, providing tools for both clinicians and individuals to monitor and manage stress more effectively.

1.6 Research Gap

Despite significant advancements in stress detection through wearable sensors, the research landscape reveals gaps, particularly in the translation of complex physiological data into actionable insights. Current methodologies often overlook the nuanced interplay between individual baseline physiological metrics and their responses to stress. Furthermore, the adaptation of stress detection models to diverse, real-world environments beyond controlled laboratory settings still needs to be explored. The need for models that account for individual variability in stress responses and can operate in dynamic, unstructured environments represent a critical research gap.

1.7 Problem Statement and Objectives

The problem at hand is the identification and prediction of stress levels in healthcare workers, using wearable sensor data that captures their physiological state in a high-stress environment. The objective is to develop an advanced machine learning framework that can accurately classify stress levels, identify the most significant physiological predictors of stress, and balance the data for improved model performance and generalization across different individuals and settings.

1.8 Chapter Organization

This thesis is meticulously structured to navigate through the intricate processes of stress detection using advanced machine learning techniques applied to wearable sensor data within healthcare settings. The organization of the chapters is as follows:

- **Chapter 1** sets the stage with an overview of the stress landscape within healthcare environments, emphasizing the critical role of wearable technology in stress detection and the contribution of machine learning techniques to this field.
- **Chapter 2** dives into a comprehensive literature review, dissecting existing research on stress detection methodologies, wearable sensors' evolution, and machine learning models' application in this domain. It also highlights the research gaps that this study aims to address.
- **Chapter 3** delineates the research methodology, detailing the data collection process from wearable sensors, data preprocessing, feature engineering, and the balancing techniques employed, SMOTE, to ensure robust model training and validation.
- **Chapter 4** delves into the machine learning techniques utilized in the study, including Random Forest (RF), Extreme Gradient Boosting (XGB), and Artificial Neural Networks (ANN). It explains the rationale behind selecting these models and their expected efficacy in the context of stress prediction.
- **Chapter 5** offers a detailed account of the implementation process, showcasing how the methodologies described in the previous chapters are applied to the dataset. It covers the practical aspects of model training, validation, and the fine-tuning of hyperparameters to optimize performance.
- **Chapter 6** presents the results and discussions emanating from the application of these models. It provides a critical analysis of the outcomes, comparing the models' performance and discussing the implications of the findings in the broader context of stress management in healthcare.

- **Chapter 7** encapsulates the study by summarizing the essential findings and offering recommendations based on the research outcomes. It also provides a roadmap for future research, suggesting potential areas for further exploration and refinement of stress detection techniques using wearable sensor data.

Each chapter builds upon the preceding one, threading a narrative that evolves from conceptual foundations to empirical validations, culminating in a set of actionable insights and recommendations for the field.

1.9 Conclusion

This chapter sets the stage for the ensuing research, outlining the pervasive issue of stress in healthcare, particularly among nurses, and the imperative for sophisticated detection methodologies. It positions the study within the context of existing literature, identifying gaps that the current research aims to fill. The chapter establishes the objectives of leveraging advanced machine learning models to interpret wearable sensor data and predict stress levels, which could inform strategies for mitigating stress in high-risk professions. The organization of the report provides a clear roadmap for the reader, encapsulating the journey from theoretical underpinnings to practical applications.

CHAPTER 2: LITERATURE REVIEW

This chapter provides a detailed review of the existing research related to stress detection using wearable sensors and machine learning techniques, specifically in the context of healthcare settings. The review encompasses various approaches and models, discussing their strengths, limitations, and the context in which they were applied.

2.1 Types of Stress Detection Techniques

This section provides a comprehensive literature review of existing research on stress detection in healthcare, focusing on various approaches and techniques.

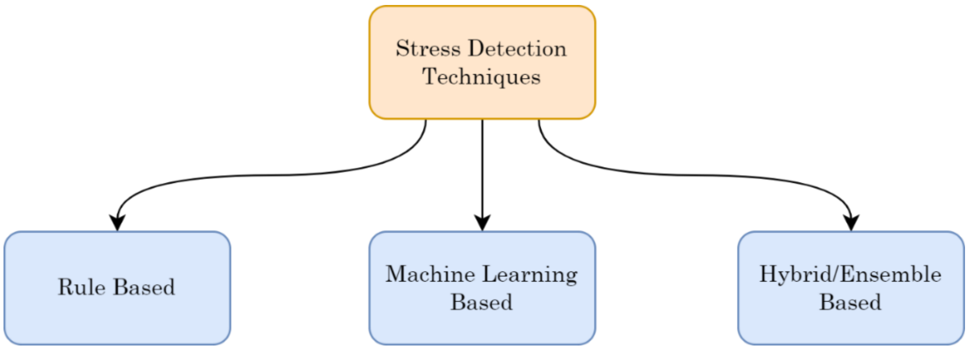


Figure 2.1: Types of Stress Detection Techniques

2.1.1 Rule-Based Approaches:

This section explores traditional stress detection methods, primarily focusing on rule-based systems. These systems rely on predefined criteria or thresholds for stress indicators, like heart rate or cortisol levels, to determine stress levels.

2.1.2 Machine Learning-based Approaches:

A discussion on the use of machine learning algorithms in stress detection. This includes an overview of different models like Random Forest, XGBoost, and ANN, emphasising their application in interpreting complex patterns in physiological data for stress prediction.

2.1.3 Hybrid/Ensemble Based Approaches:

Examination of advanced techniques combining multiple machine learning models for improved accuracy. This includes the use of ensemble methods like Stacking Classifiers, which integrate the strengths of individual models to enhance predictive performance.

2.2 Features and Parameters

A detailed discussion on the fundamental physiological and behavioural features used in stress detection, including heart rate variability, electrodermal activity, and others. This section also delves into the parameters critical for the effective functioning of the models and their impact on stress detection accuracy.

2.3 Evaluation Metrics

Overview of various metrics used to evaluate the performance of stress detection models. This includes accuracy, precision, recall, F1-score, and ROC-AUC, highlighting how these metrics are applied to assess the effectiveness of different models.

2.4 Comparison with Existing Models

A comparative analysis of the proposed models with existing stress detection models in the literature. This section discusses the advancements made over previous models, highlighting improvements in terms of accuracy, reliability, and applicability in real-world healthcare settings.

2.5 A Review of the Work Done in the Field

The following reviews provide a comprehensive survey about the developments in the stress detection technology around the world.

"Stress Monitoring Using Wearable Sensors: A Pilot Study and Stress-Predict Dataset," by **Talha Iqbal et al. [1]**, describes a study focused on stress detection using wearable sensors. It details a pilot study where 35 volunteers underwent stress-inducing tasks while wearing wrist-worn watches with photoplethysmogram (PPG) sensors. The study aimed to collect physiological data for stress monitoring and to perform statistical analyses to understand the association between various physiological variables and stress levels. Key components of the study included the use of Empatica E4 watches for data collection and novel algorithms for estimating respiratory rates. The study's findings contribute to the development of reliable, valid, and sensitive physiological stress monitoring using wearable technologies.

"Stress Classification Using Brain Signals Based on LSTM Network," by **Nishtha Phutela et al. [2]**, explores stress detection using electroencephalography (EEG) signals. It proposes a stress classification system employing an EEG signal from 35 volunteers, using a 4-electrode Muse EEG headband while viewing four movie clips with varying emotional content. The study compares Multilayer Perceptron (MLP) and Long Short-Term Memory (LSTM) models for classifying stress and non-stress groups, achieving a maximum classification accuracy of 93.17% with a two-layer LSTM architecture. This research highlights the effectiveness of LSTM networks in stress classification using EEG signals.

"Ensemble Machine Learning Model Trained on a New Synthesized Dataset Generalizes Well for Stress Prediction Using Wearable Devices," by **Gideon Vos et al. [3]**, presents an innovative approach to stress prediction using wearable sensor data. This research combines data from multiple small datasets to form a more extensive, synthesised dataset named StressData, which improves the generalisation of machine learning models. The study utilises ensemble machine learning techniques, combining gradient boosting with an artificial neural network. The findings demonstrate a significant improvement in predictive accuracy on new, unseen validation data. The approach addresses challenges in stress prediction, like dataset variability, class imbalance, and model generalisation, making a significant contribution to the field.

"Occupational Stress and Burnout Among Intensive Care Unit Nurses During the Pandemic: A Prospective Longitudinal Study of Nurses in COVID and Non-COVID Units," by **Pratima Saravanan et al. [4]** presents a comprehensive study of stress levels in nurses during the COVID-19 pandemic. This study employed wearable sensors to collect physiological stress data from ICU nurses in both COVID and non-COVID units. It integrated these measurements with self-reported stress and burnout questionnaires, providing a unique perspective on the impact of the pandemic on healthcare professionals. The study's findings highlighted the significant levels of stress and burnout experienced by nurses, particularly those in COVID-19 units. They underscored the importance of using both objective and subjective data in understanding occupational stress.

"Stress Monitoring Using Wearable Sensors: IoT Techniques in the Medical Field," by **Fatma M. Talaat, Rana Mohamed El-Balka [5]**, delves into the integration of IoT and Machine Learning (ML) technologies for stress monitoring using wearable sensors. It emphasises the potential of IoT in enhancing healthcare services by providing real-time stress monitoring and personalised feedback. The paper proposes a new Stress Monitoring Algorithm (SMA) using wearable sensors, comparing various ML algorithms for stress prediction. It demonstrates that Random Forest and Optimised Support Vector Machine (OSM) classifiers are highly effective for stress classification, with significant implications for healthcare services using IoT and wearable technology.

"Machine Learning and IoT for Prediction and Detection of Stress," by **Purnendu Shekhar Pandey [6]**, examines the application of Machine Learning (ML) and Internet of Things (IoT) technologies for stress detection. The paper emphasises the use of heart rate data as a critical parameter for stress detection. It discusses a remote stress detector, an IoT device that assesses stress levels using heartbeat readings. The paper elaborates on the methodology, hardware implementation, and classification techniques for stress detection, including Logistic Regression and Support Vector Machine (SVM). The study underscores the importance of accurate heart rate monitoring and the potential integration of such systems into health monitoring devices.

"Detection of Stress Using Wearable Sensors in IoT Platform," by **Sreedevi Uday et al. [7]**, focuses on developing an IoT-based system for stress detection. The study uses an intelligent band and chest strap module to measure electrodermal activity (EDA) and heart rate (HR), with data sent to a cloud-based ThingSpeak server. MATLAB Visualization applications compute the data to provide stress reports. The paper highlights the use of physiological parameters in real-time stress detection, emphasising the integration of IoT and wearable technology for health monitoring.

"An IoT-based Real-Time Stress Detection System for Firefighters," by **Jeril V Raj, Sarath T V [8]**, discusses a novel system for stress monitoring in high-risk professions, specifically targeting firefighters. It presents a real-time monitoring system that employs Galvanic Skin Response (GSR) and heart rate sensors embedded in gloves. The system uses an IoT platform to analyse the collected data and provide immediate feedback on the firefighters' stress levels during rescue missions. The study highlights the use of advanced technologies in hazardous work environments and emphasises the potential benefits of early stress detection and intervention.

"Smart-Pillow: An IoT-Based Device for Stress Detection Considering Sleeping Habits," by **Laavanya Rachakonda et al. [9]**, proposes an innovative IoT-based device for stress detection through sleeping habits analysis. The research uses a Smart-Pillow connected to a wireless tracker to emphasise the interrelation between stress and sleep. This device monitors various physiological parameters during sleep and classifies stress levels into five states. The study employs fuzzy logic for data analysis and emphasises the importance of understanding the relationship between sleep quality and stress, suggesting potential applications in intelligent healthcare.

"Stress Diagnostic System and Digital Medical Record Based on Internet of Things," by **Rachmad Setiawan et al. [10]**, introduces a comprehensive system for stress diagnosis incorporating IoT technology. The study focuses on designing and implementing a stress diagnostic system using fuzzy logic, employing sensors for temperature, Galvanic Skin Response (GSR), and heart rate. The data collected from these sensors is processed using fuzzy logic methods to diagnose stress levels, which are then displayed on a website and accessible through an Android application. This research highlights the use of IoT and fuzzy logic in the medical field for stress monitoring and digital medical record keeping.

"Hierarchical Deep Neural Network for Mental Stress State Detection Using IoT-Based Biomarkers," by **Akshi Kumar et al. [11]**, explores mental stress detection using a multi-level deep neural network and sensor-based bio-signals. The study employs both wrist-based and chest-based sensors, training multivariate time-series data with a hierarchy of networks to generate high-level features for stress state classification. The paper highlights the use of a model-level fusion strategy to combine these features and demonstrates its effectiveness with impressive performance accuracy. This research offers significant contributions to the field of mental stress detection using IoT and deep learning.

"A Review on Mental Stress Detection Using Wearable Sensors and Machine Learning Techniques," by **Shruti Gedam, Sanchita Paul [12]**, offers a comprehensive overview of methodologies in stress detection. It examines various wearable sensors and machine learning techniques used in stress detection, focusing on electrocardiogram (ECG), electroencephalography (EEG), and photoplethysmography (PPG) sensors. The paper evaluates stress detection in different environments, such as driving, studying, and working. It provides insights into the advantages, limitations, and issues of different studies, aiming to guide future research in this field.

"Pain and Stress Detection Using Wearable Sensors and Devices - A Review," by **Jerry Chen et al. [13]**, provides a thorough analysis of the mechanisms of pain and stress, their detection using wearable sensors, and the impact of these factors on human health. It examines various physiological and behavioural signals used in pain and stress assessment, including heart activity, brain activity, muscle activity, electrodermal activity, and more. The review also discusses various medical devices and wearable sensors used in detecting pain and stress, highlighting the advancements in IoT and wearable technology in healthcare.

"Automatic Stress Detection Using Wearable Sensors and Machine Learning: A Review," by **Shruti Gedam, Sanchita Paul [14]**, provides a detailed analysis of stress detection techniques using wearable sensors and machine learning. It covers various aspects of stress detection, including feature extraction, physiological measures, and the application of machine learning algorithms. The paper reviews several studies, highlighting the effectiveness of wearable sensing and machine learning in stress detection and offering a comprehensive guide for efficient stress detection methods.

"Personalised Stress Monitoring Using Wearable Sensors in Everyday Settings," by **Ali Tazarv et al. [15]**, investigates stress monitoring using heart rate (HR) and heart rate variability (HRV) data from photoplethysmography (PPG) sensors in everyday environments. It introduces a layered system architecture for data collection and labelling, using a smart strategy to obtain informative samples for labelling. The study involved fourteen volunteers over 1-3 months and utilised machine learning methods for binary stress detection. Preliminary results showed the detector's fair accuracy, laying the groundwork for more sophisticated, context-aware, personalised models for health professionals.

"Human Stress Detection With Wearable Sensors Using Convolutional Neural Networks," by **Manuel Gil-Martin et al. [16]**, focuses on using convolutional neural networks (CNNs) for stress detection from wearable sensor data. The study uses a combination of inertial and physiological signals processed through deep learning architectures for feature extraction and stress detection. The paper demonstrates the effectiveness of CNNs in interpreting complex biosignal data for stress monitoring, showing significant advancements in the field of stress detection using machine learning and wearable technologies.

"Stress Detection and Relief Using Wearable Physiological Sensors," by **Kriti Sethi et al. [17]**, presents a unique concept for detecting and relieving stress using a cap equipped with physiological sensors. This prototype cap uses a combination of sensors to detect stress levels based on heart rate and brainwave activity. It also incorporates immediate relief measures like auditory stimulation and scalp massage. This innovative approach emphasises early detection and intervention in stress management, integrating IoT for data recording and transmission, potentially aiding in mental health disorder prevention.

"Context-Aware Speech Stress Detection in Hospital Workers Using Bi-LSTM Classifiers," by **Amr Gaballah et al. [18]**, examines stress detection in hospital workers using bi-directional Long Short-Term Memory (Bi-LSTM) networks. It uses a context-aware approach, considering factors like location within the hospital and circadian rhythms alongside speech features. The study involved a large sample of hospital workers monitored over ten weeks. It demonstrates the significance of context in enhancing stress prediction accuracy, showing notable improvements in detection when both location and circadian rhythm data are included.

"Automated Stress Level Detection for Hospital Nurses: A Single Triaxial Wearable Accelerometer Sensor System Approach," by Mohammad Sakib, **Syeda Shanaz Pervez [19]**, focuses on stress detection in nurses using a single triaxial wearable accelerometer sensor. This study emphasises the use of machine learning methods like k-NN, Neural Networks, SVM, Naive Bayes, and Ensemble Classifiers for assessing nurses' stress levels. It demonstrates a novel approach to stress monitoring in healthcare settings, leveraging accelerometer data to provide insights into the stress levels of nurses during their shifts.

"Occupational Stress Monitoring Using Biomarkers and Smartwatches: A Systematic Review," by **Analúcia Morales et al. [20]**, provides an extensive overview of wearable technologies and biomarkers for occupational stress monitoring. This review is part of a larger project aimed at developing a stress surveillance system. It follows the PRISMA guidelines and selects 38 articles for in-depth analysis, focusing on the types of technologies and biomarkers used in the studies. The paper underscores that while stress assessments are primarily based on standardised questionnaires, recent advancements in wrist wearables, incorporating physiological and chemical sensors, show promise in stress identification.

2.6 Extract of the Literature

Table 2.1 Literature Survey based on Methodology Used and their Outcomes

S. No	Year	Author	Paper Title	Methodology Used	Outcomes (Accuracy)	Dataset Used
1	2022	Seyedmajid Hosseini, et al.	A Multimodal Sensor Dataset for Continuous Stress Detection of Nurses in a Hospital	Random Forest	The study focuses on dataset creation and contextual analysis rather than a specific accuracy metric.	Biometric data (EDA, HR, Temp) from nurses in a hospital setting

2	2022	Talha Iqbal et al.	Stress Monitoring Using Wearable Sensors: A Pilot Study and Stress-Predict Dataset	Use of Empatica E4 watches to record BVP, IBI, and heart rate, involving stress-inducing tasks.	The paper primarily focuses on developing a dataset and does not specify a particular accuracy metric.	Custom dataset consisting of BVP, IBI, and heart rate data from 35 volunteers.
3	2022	Nishtha Phutela et al.	Stress Classification Using Brain Signals Based on LSTM Network	LSTM network for classifying stress using EEG signals.	Achieved a maximum classification accuracy of 93.17% using a two-layer LSTM architecture.	EEG signals from 35 volunteers using a 4-electrode Muse EEG headband.
4	2023	Gideon Vos et al.	Ensemble Machine Learning Model Trained on a New Synthesized Dataset Generalizes Well for Stress Prediction Using Wearable Devices	Ensemble model combining gradient boosting and an artificial neural network, trained on Synthesized Stress Data.	85% predictive accuracy on new unseen validation data.	Synthesized Stress Data, a combination of several smaller datasets.
5	2023	Pratima Saravanan et al.	Occupational Stress and Burnout Among Intensive Care Unit Nurses During the Pandemic	Prospective longitudinal mixed-methods study using wrist-worn wearable sensors and validated questionnaires.	The study provides a detailed analysis of stress and burnout levels but does not specify a singular accuracy metric.	Data from ICU nurses working in COVID and non-COVID units.
6	2023	Fatma M. Talaat, Rana Mohamed El-Balka	Stress Monitoring Using Wearable Sensors: IoT Techniques in Medical Field	Development of SMA using wearable sensors integrated with IoT and ML.	Best performance achieved by Random Forest and OSM.	2001 samples including humidity, temperature, step count, and stress levels.
7	2017	Purnendu Shekhar Pandey	Machine Learning and IoT for Prediction and Detection of Stress	Integration of ML (Logistic Regression, SVM) and IoT for stress detection using heart rate data.	66% for Logistic Regression, 68% for SVM.	Heart rate data, details unspecified in the excerpt.
8	2018	Sreedevi Uday et al.	Detection of Stress using Wearable Sensors in IoT Platform	Real-time monitoring of EDA and HR using wearable sensors, data transmission to ThingSpeak, and MATLAB Visualization for analysis.	The paper focuses on system development and continuous monitoring, rather than a specific accuracy metric.	Real-time EDA and HR data from the study's subjects.
9	2019	Jeril V Raj, Sarath T V	An IoT based Real-Time Stress Detection System for Fire-Fighters	Real-time stress monitoring using GSR and heart rate sensors in	The paper emphasizes system design and real-time	Real-time stress data from firefighters

				firefighter gloves, IoT communication via MQTT.	monitoring capabilities rather than a specific accuracy metric.	during rescue missions.
10	2023	Laavanya Rachakonda et al.	Smart-Pillow: An IoT Based Device for Stress Detection Considering Sleeping Habits	Monitoring and analyzing physiological and non-physiological sleeping parameters using Smart-Pillow.	The paper focuses on the development of the device and does not specify a particular accuracy metric.	Data based on sleeping habits including physiological and non-physiological parameters.
11	2019	Rachmad Setiawan et al.	Stress Diagnostic System and Digital Medical Record Based on Internet of Things	IoT system with temperature, GSR, and heart rate sensors; fuzzy logic for stress diagnosis.	Data transmission and storage rate of 100%, system working rate of 80%.	Data from 15 subjects tested five times each, including temperature, skin conductance, and heart rate.
12	2021	Akshi Kumar et al.	Hierarchical Deep Neural Network for Mental Stress State Detection Using IoT Based Biomarkers	Hierarchical deep neural network using wrist-based and chest-based sensor bio-signals.	87.7% accuracy on the WESAD benchmark dataset.	WESAD benchmark dataset with wrist and chest sensor bio-signals.
13	2021	Shruti Gedam, Sanchita Paul	A Review on Mental Stress Detection Using Wearable Sensors and Machine Learning Techniques	Comprehensive review of various stress detection methods, wearable sensors, and machine learning techniques.	Not applicable, as it's a review paper.	Not applicable, as it's a review paper.
14	2021	Jerry Chen et al.	Pain and Stress Detection Using Wearable Sensors and Devices - A Review	Review of mechanisms, classifications, detection methods, and wearable sensors for pain and stress detection.	Not applicable, as it's a review paper.	Not applicable, as it's a review paper.
15	2020	Shruti Gedam, Sanchita Paul	Automatic Stress Detection Using Wearable Sensors and Machine Learning: A Review	Review of studies on wearable sensors and machine learning for stress detection.	Not applicable, as it's a review paper.	Not applicable, as it's a review paper.
16	2021	Ali Tazarv et al.	Personalized Stress Monitoring using Wearable Sensors in Everyday Settings	Use of PPG sensors in smartwatches and machine learning algorithms for stress detection.	Macro-F1 score of up to 76% for binary stress classification.	Data from 14 volunteers over 1-3 months, involving heart rate and heart rate variability.
17	2022	Manuel Gil-Martin et al.	Human Stress Detection With Wearable Sensors Using Convolutional Neural Networks	Deep learning architecture with CNNs, using physiological signals from wearable devices.	High performance in stress detection using LOSO cross-validation.	WESAD dataset.

18	2019	Kriti Sethi et al.	Stress Detection and Relief Using Wearable Physiological Sensors	Development of a cap with PPG and EEG sensors for stress detection and relief measures.	Focuses on stress detection and relief rather than specific accuracy.	Physiological data collected using PPG and EEG sensors.
19	2021	Amr Gaballah et al.	Context-Aware Speech Stress Detection in Hospital Workers Using Bi-LSTM Classifiers	Bi-LSTM neural network for speech analysis combined with context-aware data (location and circadian rhythm).	Improvement in stress detection accuracy and F1 scores with the inclusion of context.	Speech and contextual data from 144 hospital workers over 10 weeks.
20	2023	Mohammad Sakib, Syeda Shanaz Pervez	Automated Stress Level Detection for Hospital Nurses: A Single Triaxial Wearable Accelerometer Sensor System Approach	Use of a single triaxial wearable accelerometer sensor; multiple machine learning methods for stress detection.	Best average performance of 80.4% accuracy by Medium Gaussian-SVM.	Accelerometer data from hospital nurses.
21	2022	Analúcia Morales et al.	Occupational Stress Monitoring Using Biomarkers and Smartwatches: A Systematic Review	Systematic review of wrist wearables and biomarkers for stress detection.	Not applicable, as it's a review paper.	Not applicable, as it's a review paper.

2.7 Conclusion

This chapter has provided a comprehensive review of the current literature on stress detection, specifically focusing on the application of wearable sensor data and machine learning techniques in healthcare settings. The review delved into various approaches for stress detection, including rule-based, machine learning-based, and hybrid/ensemble methods. It highlighted the evolution of stress detection techniques from traditional heuristic methods to more sophisticated, data-driven approaches.

Through the examination of different methodologies, this review emphasized the increasing role of machine learning in interpreting complex physiological data for stress prediction. It explored how different models such as Random Forest, XGBoost, and Artificial Neural Networks have been applied to detect stress, discussing their respective strengths and limitations. The review also underlined the significance of hybrid or ensemble approaches, which synergize multiple algorithms to enhance predictive accuracy.

Additionally, the review shed light on the crucial features and parameters that are instrumental in stress detection, such as heart rate variability and electrodermal activity. It discussed how the careful selection and analysis of these features are vital for the accuracy and effectiveness of stress detection models. The evaluation metrics used to assess these models, including accuracy, precision, recall, and F1-score, were also examined, providing a framework for comparing and understanding the performance of various approaches.

In summary, this chapter has laid the groundwork for understanding the complex landscape of stress detection in healthcare. It has set the stage for the subsequent chapters, which will delve into the methodology of applying these techniques to real-world data, aiming to contribute significantly to the field of stress management and healthcare technology. The following chapter will detail the methods employed in this study, encompassing data collection, feature selection, machine learning algorithm application, and model evaluation, thereby continuing the exploration into advanced stress detection and monitoring.

CHAPTER 3: RESEARCH METHODOLOGY

This chapter delineates the comprehensive research methodology employed in this study, focused on optimizing stress prediction using wearable sensor data through advanced machine learning techniques. It details the systematic process of data collection, pre-processing, feature engineering, implementation of machine learning models, and their evaluation. This methodology is designed to address the nuances of stress detection in a healthcare context, leveraging the potential of wearable technologies and sophisticated data analysis techniques. The chapter commences with a discussion on the research design, followed by an elaboration on the experimental setup, proposed methodology, data collection, and pre-processing. It concludes with the model implementation and evaluation.

3.1 Research Design

The study adopts an empirical research design, incorporating a mix of descriptive and exploratory methods aimed at systematically investigating stress detection through wearable sensors. The research is structured to not only describe the characteristics of stress data obtained from wearable devices but also to explore the potential of various machine learning techniques in accurately predicting stress levels. This comparative approach facilitates the evaluation of different models, offering insights into their efficiency and applicability in real-world healthcare scenarios.

The research design revolves around several key components:

1. **Collection and Analysis of Wearable Sensor Data:** Gathering extensive physiological data from wearable devices, which includes metrics such as heart rate, electrodermal activity (EDA), and skin temperature, to form a comprehensive dataset for analysis.
2. **Feature Engineering:** Refining the dataset by engineering features that are crucial in stress detection. This involves identifying and extracting relevant features that significantly contribute to the accuracy of stress prediction.
3. **Implementation of Machine Learning Models:** Applying a range of machine learning models, including Random Forest, XGBoost, Artificial Neural Networks (ANN), and ensemble methods, to the processed data. This step is crucial in understanding how different algorithms interpret and predict stress based on physiological data.
4. **Evaluation of Model Performance:** Applying a range of machine learning models, including Random Forest, XGBoost, Artificial Neural Networks (ANN), and ensemble methods, to the processed data. This step is crucial in understanding how different algorithms interpret and predict stress based on physiological data.

5. **Comparative Analysis:** Conducting a comparative analysis of the models to identify the most effective approach for stress detection using wearable sensor data. This includes comparing the performance of individual models against hybrid or ensemble approaches.

The research design is tailored to address the complexities of stress detection in healthcare, focusing on the potential of wearable sensors as a non-invasive and continuous monitoring tool. By exploring different machine learning strategies, the study aims to advance the field of stress detection, contributing valuable insights into the development of effective stress management solutions in healthcare.

3.2 Experimental Setup

The experimental setup for this study was meticulously designed to ensure the accurate and efficient application of machine learning models for stress detection using wearable sensor data. The setup involved the following components:

1. **Development Environment**:
   Programming Platform: Python, known for its extensive libraries and community support for data analysis and machine learning, was the primary programming language used.
   Analytical Tools: Jupyter Notebooks provided an interactive environment for code execution, visualization, and iterative testing.
2. **Data Collection and Preparation:**
   Data was sourced from wearable sensors, encompassing metrics such as heart rate, electrodermal activity (EDA), and skin temperature.
   The dataset, comprising approximately 11.5 million entries, included multidimensional data reflecting physiological and orientation measurements.
3. **Machine Learning Framework:**
   Utilization of various libraries like pandas for data manipulation, NumPy for numerical operations, Seaborn and matplotlib for data visualization, and scikit-learn, along with imblearn for machine learning model implementation and evaluation.
4. **Modelling Process:**
   The methodology involved data pre-processing (handling missing values, feature scaling), and applying techniques like SMOTE for addressing class imbalance.
   The stress classification models included RandomForest, XGBoost, MLPClassifier, and a Stacking Classifier for ensemble learning.
5. **Validation Technique:**
   Train-test split and cross-validation methods are employed to evaluate the models' performance and generalizability on unseen data.
6. **Performance Metrics:**

Key metrics such as accuracy, precision, recall, F1-score, and ROC-AUC were used for evaluating model performance.

3.3 Proposed Methodology

The proposed methodology in this study integrated a systematic approach to stress detection using machine learning models applied to wearable sensor data. The following steps outline the methodology:

1. **Data Loading and Preliminary Analysis:**
   Loading the dataset into a Python environment using pandas.
   Conducting an initial statistical analysis to understand data distributions and characteristics.
2. **Feature Engineering:**
   Extracting relevant features from the dataset, including physiological signals (heart rate, EDA, temperature) and orientation data (X, Y, Z axes).
   Application of statistical methods (like skewness and kurtosis) to derive additional insights from the data.
3. **Data Preprocessing:**
   Addressing missing values and anomalies in the dataset.
   Standardizing features using techniques like StandardScaler for consistent model training.
4. **Handling Imbalanced Data:**
   Implementing Random UnderSampling and SMOTE to balance the dataset, ensuring an equitable representation of different stress levels.
5. **Model Development and Training:**
   Training various models such as RandomForestClassifier, XGBClassifier, and MLPClassifier.
   Developing an ensemble model using StackingClassifier, combining the strengths of individual models for improved performance.
6. **Model Evaluation:**
   Assessing each model's performance using metrics like accuracy, ROC-AUC, and generating classification reports.
   Analyzing confusion matrices to understand model predictions in detail.
7. **Feature Importance Analysis:**
   Utilizing permutation is essential to evaluate the impact of different features on the model's predictive ability.
8. **Final Model Selection:**
   Comparing the performance of all models to select the most effective approach for stress detection.

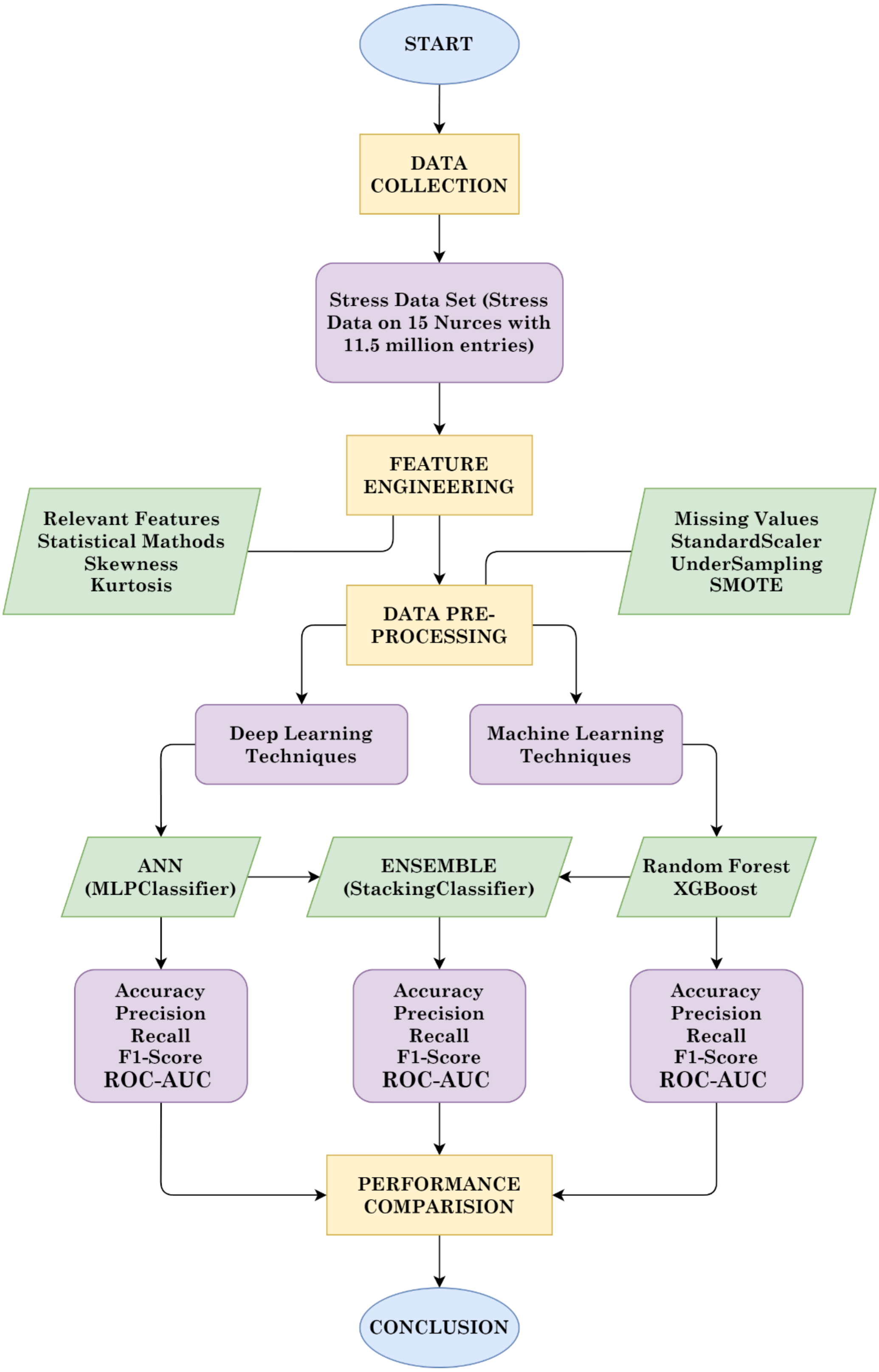


Figure 3.1: Proposed Model

his methodology embodies a comprehensive approach, from data pre-processing to model evaluation, ensuring a thorough investigation into the effectiveness of machine learning techniques in stress detection using wearable sensor data.

3.4 Data Collection

Data collection in this study was focused on acquiring extensive and accurate physiological data from wearable sensors, essential for stress prediction. The data collection process involved several key steps:

1. **Source of Data:**
   The dataset was specifically chosen for its relevance in continuous stress monitoring, particularly among healthcare professionals during the COVID-19 pandemic.
   Wearable devices, such as fitness trackers and smartwatches, were used to collect data, focusing on metrics like heart rate, electrodermal activity (EDA), and skin temperature.
2. **Data Acquisition:**
   Data was gathered over a defined period, ensuring a substantial dataset for analysis.
   The dataset included a diverse range of physiological signals, offering a holistic view of the subjects' stress levels.
3. **Data Integrity and Confidentiality:**
   Strict measures were employed to ensure the integrity and confidentiality of the data, adhering to ethical standards and privacy regulations.

3.4.1 Data Set

The dataset used in this study was curated explicitly for stress detection, comprising approximately 11.5 million entries with multiple features. Critical aspects of the dataset included:

1. **Data Attributes:** The dataset contained a range of physiological and orientation data, including heart rate (HR), electrodermal activity (EDA), skin temperature (TEMP), and orientation data (X, Y, Z axes).
2. **Data Structure:** Data was organized in a structured format, suitable for processing and analysis in machine learning models.
3. **Sample Size and Diversity:** The dataset's large size and the diversity of its entries made it a robust source for training and testing the machine learning models.

The data in the dataset is represented in the CSV (Comma Separated Values) format with the following attributes: X, Y, Z, EDA, HR, TEMP, id, datetime, and label.

Table 3.1: Properties of Dataset

Aspect	Detail
Data Set Characteristics	Text
Number of Instances	11,509,051
Area	Health

Attribute Characteristics	N/A
Number of Attributes	9
Date Collected	2023-01-20
Associated Tasks	Classification
Missing Values	No
Number of Web Hits	N/A

Table 3.2: Attributes of Dataset

Feature	Description	Unique Entries
X	Orientation data (X-axis)	256
Y	Orientation data (Y-axis)	256
Z	Orientation data (Z-axis)	256
EDA	Electrodermal Activity	274,452
HR	Heart Rate	6,268
TEMP	Temperature	599
id	Identifier	18 Nurses
datetime	Date and Time	Approx. 10.6 million
label	Stress Label	3 categories (0,1,2)

3.5 Data Pre-processing

Data pre-processing was a critical step in preparing the dataset for machine learning model implementation. The process involved:

1. **Cleaning and Formatting:**
   Removing any irrelevant or redundant data, ensuring the dataset was focused and concise.
   Formatting the data to a uniform structure, facilitating easier manipulation and analysis.
2. **Handling Missing Values:**
   Identifying and addressing missing values in the dataset, either by imputation or removal, to maintain data quality.
3. **Normalization and Standardization:**
   Normalizing data to bring all the features to a similar scale, thus avoiding any bias due to the variance in measurement scales.
   Standardizing the dataset using methods like Z-score normalization to ensure that outliers or anomalies did not skew the model.
4. **Feature Selection:**
   Identifying and selecting the most relevant features for stress detection based on their significance and impact on model performance.

Employing techniques such as correlation analysis to ascertain the relationship between different features and the target variable.

5. **Data Transformation:**
   Converting the data into a format suitable for machine learning models, including encoding categorical variables and creating dummy variables where necessary.

This pre-processing stage was essential in enhancing the quality and usability of the data, ensuring that the machine learning models had a solid foundation for training and prediction.

3.6 Feature Importance

Feature importance in this study was integral to understanding the impact of various physiological and orientation metrics on stress prediction. This step involved:

1. **Identifying Key Features:**
   Analyzing the dataset to determine which features most significantly influence stress levels.
   Using statistical methods and machine learning algorithms to quantify the importance of each feature.
2. **Permutation Importance:**
   Employing permutation importance techniques to evaluate how the model's performance changes when each feature's values are randomly shuffled.
   This method offers insight into which features the model relies on most for making predictions.
3. **Visual Representation:**
   Creating visualizations, such as bar charts, to depict the importance of each feature clearly. This helps in effectively communicating the results to both technical and non-technical stakeholders.

3.7 Stress Classification

Stress classification in this research was executed using several advanced machine learning models, each with its unique approach to handling the dataset:

1. **Model Implementation:**
   Implementing various models such as RandomForestClassifier, XGBClassifier, MLPClassifier, and StackingClassifier.
   Tuning model parameters to optimize their performance for the specific nature of the stress data.
2. **Multiclass Classification:**
   Addressing the challenge of classifying stress levels into multiple categories, each representing a different stress intensity.
3. **Evaluation:**
   Utilizing a range of metrics like accuracy, precision, recall, F1-score, and ROC-AUC to evaluate the performance of each classification model.

Analyzing confusion matrices to gain deeper insights into the classification results, mainly focusing on the instances of misclassification.

3.8 Conclusion

In conclusion, this chapter has outlined the comprehensive research methodology used in the study. Beginning with the collection of a rich dataset from wearable sensors, the methodology encompassed thorough data pre-processing, feature engineering, and the implementation of various machine learning models for stress classification.

Feature importance analysis played a crucial role in identifying the most influential factors in stress prediction. By evaluating and visualizing the significance of different features, the study was able to focus on the most impactful data aspects.

The classification models were carefully chosen and applied to address the multiclass nature of stress levels. Each model's performance was rigorously evaluated using standard metrics, ensuring the selection of the most effective model for stress prediction.

This methodology underscores the study's commitment to leveraging advanced machine learning techniques to improve stress detection in healthcare settings. The following chapters will delve into the results and discussions, providing insights into the effectiveness of the proposed approaches and their potential implications in the field of health monitoring and stress management.

CHAPTER 4: MACHINE LEARNING TECHNIQUES USED FOR STRESS DETECTION

In the pursuit of understanding and mitigating stress within healthcare environments, particularly among nursing staff, machine learning (ML) stands as a critical technological ally. This chapter delves into the ML techniques that underpin the analysis of physiological data derived from wearable sensors. These techniques enable the discernment of stress patterns, offering a quantitative approach to what is inherently a qualitative experience. The methodologies herein described lay the groundwork for a sophisticated detection system that not only identifies stress signals but also provides actionable insights into the overall well-being of healthcare professionals.

4.1 Overview of Basic ML Techniques

The landscape of ML is vast and varied, encompassing a plethora of algorithms, each suited to specific types of data and analysis. The Random Forest algorithm, renowned for its versatility and robustness, is at the heart of our investigation. This ensemble learning method combines the simplicity of decision trees with the power of aggregation, making it particularly adept at handling the multifaceted nature of stress-related data. The following sections introduce the fundamental ML algorithms employed in this study, providing a foundation for the following complex analyses.

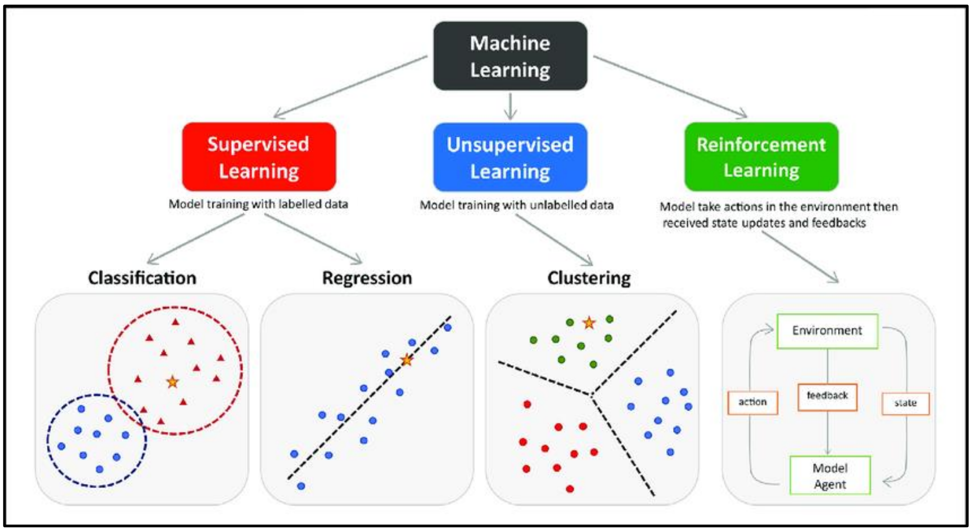


Figure 4.1: Types of Machine Learning Algorithms

4.1.1 Random Forest

Random Forest stands as a paragon of ensemble techniques, operating by constructing a multitude of decision trees during the training phase. Its ability to output the mode of the classes (classification) or mean prediction (regression) of the individual trees renders it both powerful and precise. Random Forest's inherent design mitigates the risk of overfitting, making it an excellent choice for our dataset.

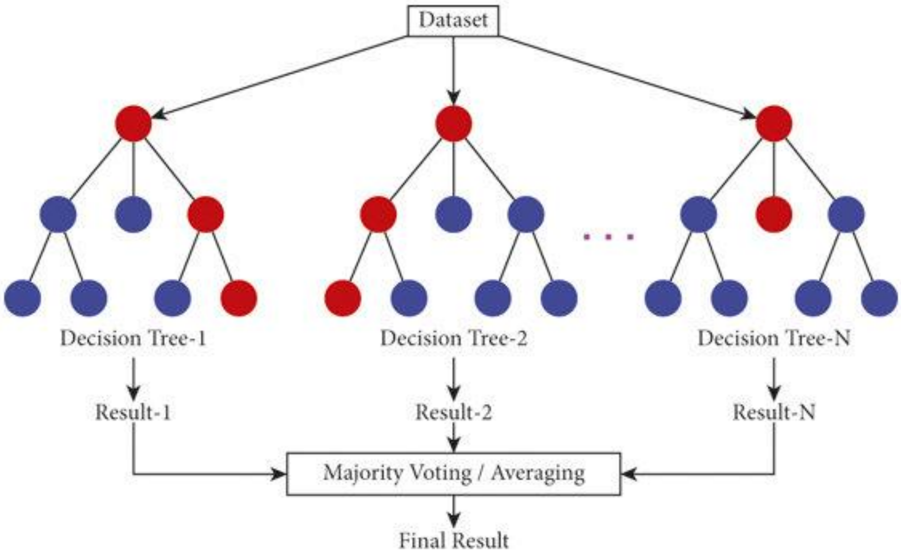


Figure 4.2 Random Forest Classifier

4.1.1.1 Applications of Random Forest

Random Forest has a broad array of applications, from medical diagnoses to stock market predictions. In the context of healthcare, it is particularly beneficial for its interpretability and the lack of need for feature scaling, allowing clinicians to understand and trust the model's predictions.

4.1.1.2 Formula for Random Forest

While Random Forest does not rely on a singular formula in the way that simpler algorithms might, its essence can be distilled into the following conceptual framework:

$$RF_{prediction} = \frac{1}{N}\sum_{i=1}^{N} DT_i(X)$$

Where $RF_{prediction}$ is the prediction of the Random Forest model, N is the number of trees, DT_i is the ith decision tree, and X is the input vector.

Note: The above is a conceptual representation, not a direct formula used in computations.

4.1.2 XGBoost

XGBoost, which stands for eXtreme Gradient Boosting, is an optimized distributed gradient boosting library designed to be highly efficient, flexible, and portable. It implements machine learning algorithms under the Gradient Boosting framework. XGBoost provides parallel tree boosting (also known as GBDT, GBM) that solves many data science problems quickly and accurately. The same code runs on a central distributed environment (Hadoop, SGE, MPI) and can solve problems beyond billions of examples.

One of the core features of XGBoost is its scalability, which drives its widespread adoption in machine learning competitions and practical applications. It is capable of performing the tasks of classification, regression, ranking, and user-defined prediction problems.

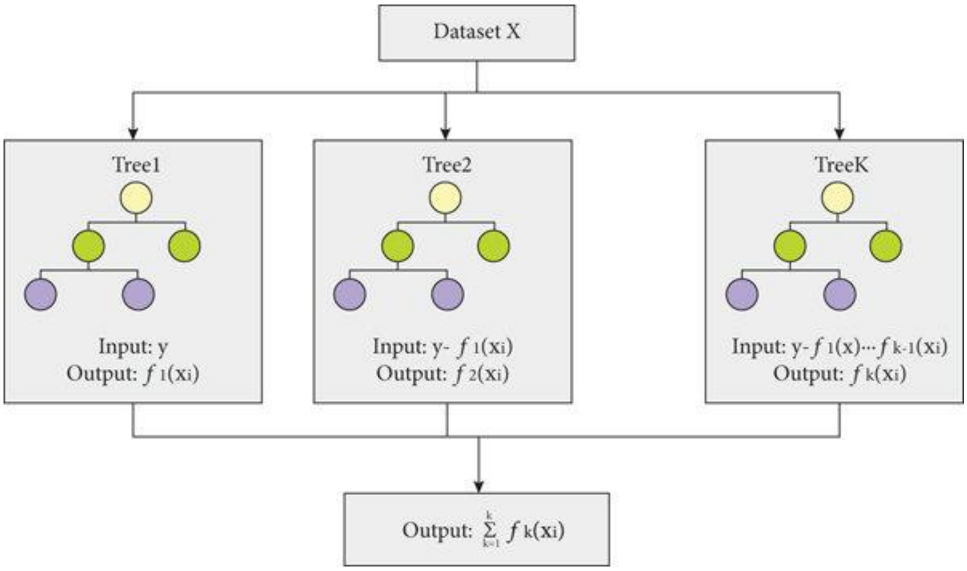


Figure 4.3: XGB Classifier

4.1.2.1 Applications of XGBoost

XGBoost has been successfully applied to a vast array of problems. Its applications include, but are not limited to:

1. Classification Problems: It determines whether an instance belongs to a particular class. This can be binary classification as well as multi-class classification.
2. Regression Problems: XGBoost can predict continuous variables, such as predicting house prices or sales forecasting.
3. Ranking Problems: XGBoost can be used for ranking tasks, such as search engines ranking the relevance of documents to a search query.
4. Anomaly Detection: It can be used in scenarios like fraud detection, where the algorithm can learn to detect transactions that do not follow the typical patterns.

4.1.2.2 Formula for XGBoost

The core idea of XGBoost is to iteratively add trees that predict the residuals or errors of prior trees merged together with a learning rate. The prediction model of XGBoost at iteration t can be represented as:

$$\hat{y}(x)_t = \hat{y}(x)_{(t-1)} + eta \times f_t(x)$$

Where:

- $\hat{y}(x)_t$ is the prediction at iteration t.
- $\hat{y}(x)_{(t-1)}$ is the prediction from the previous iteration.
- eta is the learning rate, which scales the contribution of each new tree.
- $f_t(x)$ is the prediction of the new tree at iteration t.

Each new tree $f_t(x)$ is fitted on the negative gradient of the loss function regarding the prediction of the previous iteration. The objective function is minimized to find the optimal tree structure, including regularization terms that penalize the complexity of the model to prevent overfitting.

4.2 Overview of Deep Learning Techniques

Deep Learning (DL) is a subset of machine learning where artificial neural networks and algorithms inspired by the human brain learn from large amounts of data. Similar to how we learn from experience, the deep learning algorithm would perform a task repeatedly, each time tweaking it a little to improve the outcome. Deep learning techniques are potent for identifying patterns in unstructured data such as images, sound, text, and time-series data.

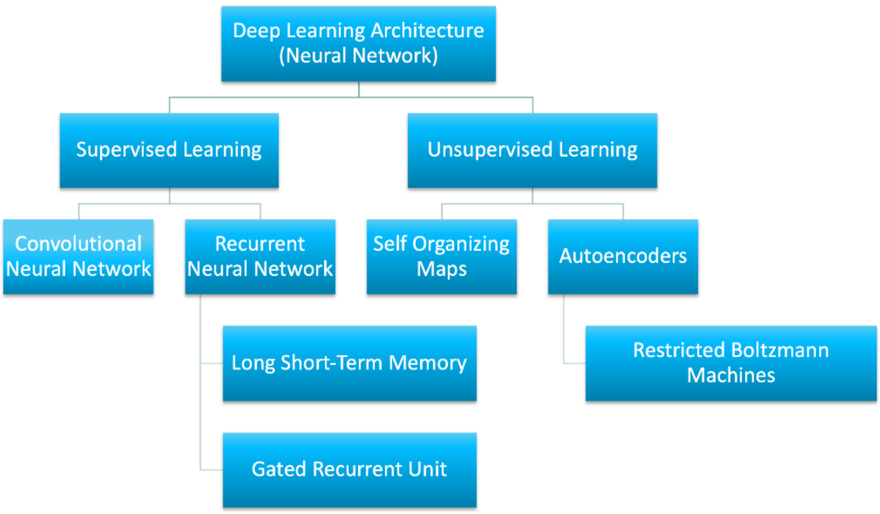


Figure 4.4: Types of Deep Learning Algorithms

4.2.1 Artificial Neural Networks (ANN)

ANNs are the foundation of deep learning. An ANN is configured for a specific application, such as pattern recognition or data classification, through a learning process. It consists of an input layer, several hidden layers, and an output layer. Each layer comprises units (neurons) that transform the input data into something the output layer can use as a base for the final prediction.

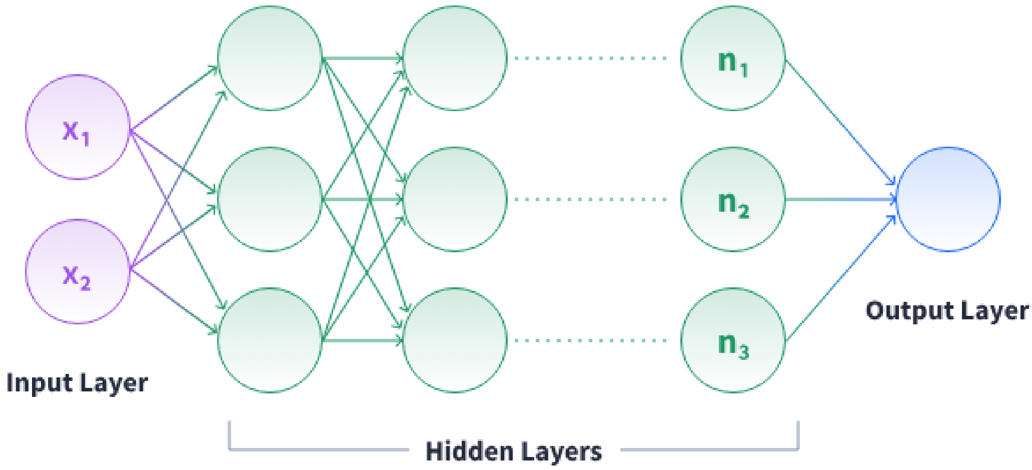


Figure 4.5: Architecture of Neural Networks

4.2.1.1 Multilayer Perceptron (MLP) Classifier

MLP is a class of feedforward artificial neural networks. An MLP consists of at least three layers of nodes: an input layer, a hidden layer, and an output layer. Except for the input nodes, each node is a neuron that uses a nonlinear activation function. MLP utilizes a supervised learning technique called backpropagation for training.

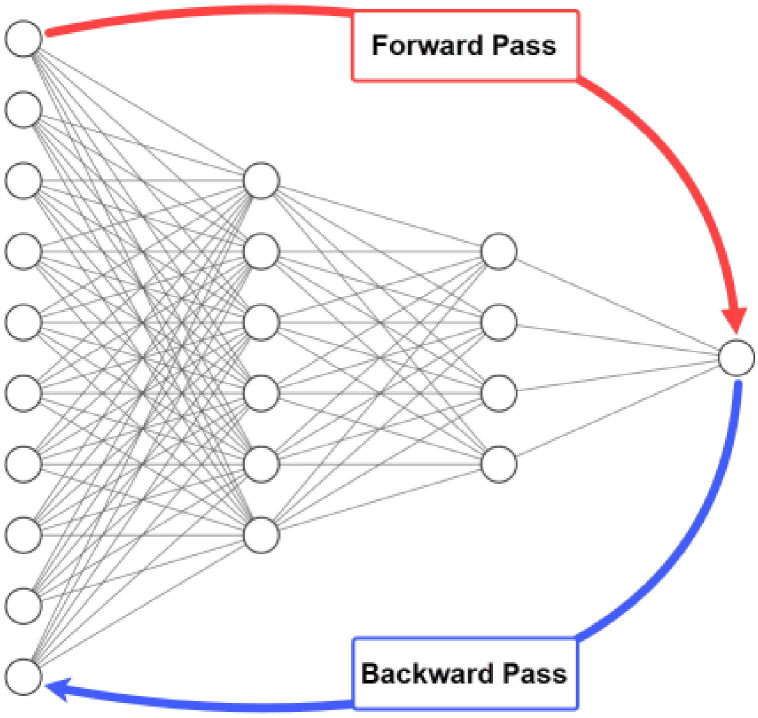


Figure 4.6: MLP Classifier

4.2.1.2 Applications of ANN/MLP

ANNs and MLPs are widely used in various applications, such as:

- Image and voice recognition
- Social network filtering
- Email spam filters
- Medical diagnosis
- Financial fraud detection
- Customer research
- Data validation
- Risk management

4.2.1.3 Formula for ANN/MLP

The general formula for a neuron's output in an ANN or MLP, which employs a weighted sum of inputs and a bias term followed by an activation function, can be expressed as:

$$O_j = \text{activation}(w_{ij} \times I_i + \text{bias}_j)$$

Where:

- O_j is the output of the neuron j.
- activation() is the activation function.
- w_{ij} is the weight from input i to neuron j.
- I_i is the input i.
- bias_j is the bias term for neuron j.

4.3 Ensemble Approach

The ensemble approach in machine learning leverages multiple learning algorithms to obtain better predictive performance than could be obtained from any of the constituent learning algorithms alone. It is a machine learning paradigm where multiple models (often called "weak learners") are trained to solve the same problem and combined to get better results. The central premise is that we can obtain more accurate and robust models when weak models are correctly combined.

	Bagging	Boosting	Stacking
Purpose	Reduce Variance	Reduce Bias	Improve Accuracy
Base Learner Types	Homogeneous	Homogeneous	Heterogeneous
Base Learner Training	Parallel	Sequential	Meta Model
Aggregation	Max Voting, Averaging	Weighted Averaging	Weighted Averaging

Figure 4.7: Types of Ensemble Approach

4.3.1 Stacking

Stacking is an ensemble learning technique that combines multiple classification or regression models via a meta-classifier or a meta-regressor. The base-level models are trained based on a complete training set; then the meta-model is trained on the outputs of the base-level model as features.

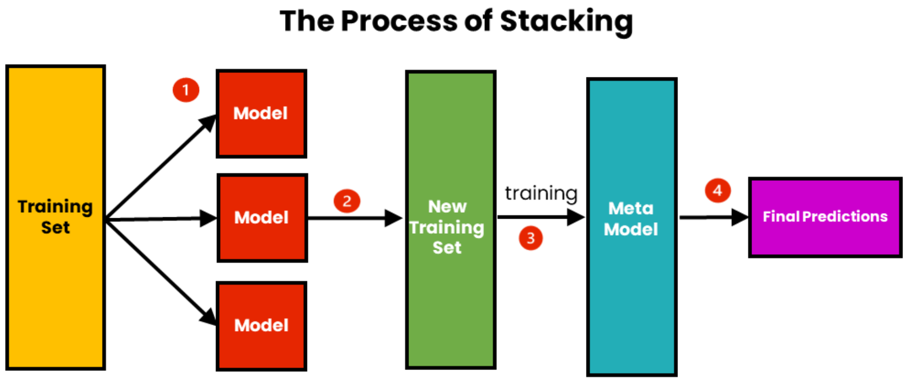


Figure 4.8: Stacking Classifier

4.3.1.1 Applications of Stacking

Stacking can be applied to a variety of fields where robust predictive performance is required, such as:

- Predictive modelling in finance for credit scoring or stock market prediction.
- Combining predictions from different models can lead to more accurate disease identification for medical diagnosis.
- Advanced analytics in sports, for instance, to predict outcomes of matches by combining various models.
- In marketing analytics, for more accurate customer segmentation and targeting.

4.3.1.2 Formula for Stacking

Stacking does not have a 'formula' in the traditional sense as it is a meta-algorithm, but the general principle can be represented as follows:

$$Meta - modelOutput = f(Basemodel_1Output, Basemodel_2Output, \ldots, Basemodel_nOutput)$$

Here, f represents the learning algorithm of the meta-model, which is trained on the outputs of the base models.

4.3.2 Hybrid Ensemble Approach

The hybrid ensemble approach combines different types of machine learning models and strategies to balance out their individual weaknesses. For example, it might combine the variance reduction advantage of bagging with the bias reduction capability of boosting.

4.4 Conclusion

This chapter has explored the spectrum of machine learning techniques from basic classifiers like Random Forest and XGBoost to complex deep learning networks such as ANNs and RNNs. It has discussed the ensemble methods that blend predictions from multiple models to enhance the robustness and accuracy of predictions, which are crucial for high-stakes applications like stress detection in healthcare.

Each method and technique comes with its own set of strengths and considerations, and the choice of algorithm often depends on the specific characteristics of the dataset and the problem at hand. The ensemble methods, with their ability to integrate diverse models, stand out for their capacity to create balanced and nuanced predictive systems.

By leveraging the hybrid ensemble approach, this research aims to pioneer in the domain of stress detection, potentially leading to more resilient and dependable models that can assist healthcare professionals in early diagnosis and intervention, thereby improving patient outcomes. The methodologies outlined here are not just theoretical constructs but are grounded in practical applications, demonstrating their viability in real-world scenarios.

Moving forward, the models and techniques presented in this chapter will be rigorously evaluated to determine their effectiveness in the context of stress detection, setting the stage for the next phase of this research endeavour.

CHAPTER 5: IMPLEMENTATION

In this chapter, we articulate the practical implementation of the methodologies and techniques that were previously outlined, with a focus on the application of machine learning and deep learning models for stress detection in healthcare professionals. The process is grounded in analysing a rich dataset, which encompasses physiological signals collected from nurses during their shifts using wearable technology. The following sections will delve into the dataset's intricacies, detailing the specifics of the data collection and the subsequent steps taken to apply various analytical models.

5.1 Dataset Description

The dataset at the heart of this research presents a unique compilation of physiological and contextual data points captured amidst the demanding environment of a hospital during the peak of the COVID-19 crisis. The data was collected in a real-world setting, reflecting genuine stress responses in a professional healthcare context.

The dataset was carefully gathered over a week from 15 dedicated female nurses within the age bracket of 30 to 55 who were at the frontline in regular hospital shifts. The collection spanned across two distinct phases – Phase-I, which stretched from April 15, 2020, to August 6, 2020, and Phase-II, which continued from October 8, 2020, to December 11, 2020. Stringent inclusion criteria were set to ensure data quality, excluding individuals with conditions that could potentially skew the stress indicators, such as pregnancy, heavy smoking, mental disorders, and chronic or cardiovascular diseases.

Electrodermal activity, heart rate, and skin temperature were at the forefront of the physiological variables, which were monitored continuously to track the nurses' stress levels. Complementing these were surveys conducted periodically via smartphones, which gathered contextual information that could correlate with the physiological indicators during stress events. The technology employed, Empatica E4, was pivotal in accurately collecting Galvanic Skin Response and Blood Volume Pulse readings.

The study was conducted following ethical guidelines, with the protocol gaining approval from the University's Institutional Review Board (FA19–50 INFOR). Nurses voluntarily participated after comprehending the study's scope and objectives, with no incentives provided to maintain impartiality. The design encompassed three phases, each involving a group of seven nurses, ensuring the data's anonymity and integrity throughout the process.

In a stride towards contributing to broader research efforts, the complete dataset, inclusive of the signals, stress events, and survey responses, has been made publicly available on Dryad, subject to privacy preservation through anonymisation techniques. Each participant was assigned a unique identifier, stripping any personal identifying information, thus safeguarding participant confidentiality.

The dataset, with its approximately 11.5 million entries, is methodically structured into nine columns, presenting a blend of continuous physiological measurements and categorical data. Orientation data is encapsulated within X, Y, and Z columns, each with 256 unique entries. Electrodermal activity is documented through 274,452 unique entries, heart rate through 6,268 unique entries, and temperature through 599 unique entries. The dataset is further enriched with categorical identifiers and extensive date and time entries, leading to a comprehensive array of data points suitable for robust stress detection analysis.

The inclusion of such diverse and continuous physiological data, coupled with periodic subjective stress evaluations, makes this dataset an unprecedented resource for advancing stress detection research. The subsequent sections will explore the application of cutting-edge machine learning models to this dataset, aiming to glean insights into the most salient physiological markers of stress and to develop an effective, reliable model for stress detection using wearable technology.

5.2 Implementation of Machine Learning Models

With the dataset meticulously described, we now transition to the application of machine learning (ML) models to identify and analyse stress indicators within the healthcare setting. This step is pivotal in translating raw data into actionable insights, enabling the development of a real-time model that can accurately detect stress levels.

The initial phase of implementing machine learning models involves preprocessing the dataset to ensure it is fit for analysis. This includes the management of any missing values, which were addressed by substituting them with median values to maintain the integrity of the dataset without introducing significant bias. Feature scaling was another critical preprocessing step to normalise the dataset. StandardScaler from the sklearn.preprocessing module was employed to transform features into a standard scale without distorting differences in the ranges of values.

The core of our implementation involved selecting robust machine learning algorithms that could handle the high dimensionality and complexity of the dataset. The RandomForestClassifier from the sklearn.ensemble module was chosen for its efficacy in handling large datasets and its ability to provide importance scores for different features. GradientBoostingClassifier, also from sklearn.ensemble, was selected for its predictive power and performance with unbalanced datasets.

The models were trained on a subset of the data, with the training set consisting of 70% of the dataset and the remaining 30% dedicated to testing. This split was conducted using the train_test_split function, ensuring that both subsets represented the whole.

Post-training, the models underwent a rigorous evaluation. Accuracy_score and confusion_matrix from the sklearn.metrics suite provided an initial assessment of performance. To delve deeper into the models' ability to classify stress correctly, a classification_report was generated, offering a breakdown of precision, recall, and F1-score for each class of stress level. The importance of each feature in predicting stress was gauged using permutation_importance, providing a clear indication of the physiological signals most correlated with stress levels.

One of the challenges faced during model training was the class imbalance within the dataset. To address this, SMOTE (Synthetic Minority Over-sampling Technique) was utilised to balance the dataset artificially, enhancing the models' ability to detect the minority class – a crucial aspect when identifying stress events that may not occur as frequently as non-stress events.

Cross-validation was performed to ensure the models were not overfitting and could generalise well to new data. This process involved splitting the training data into several smaller sets, training the models on these subsets, and validating them against complementary subsets. Hyperparameter tuning was also executed using techniques such as GridSearchCV to fine-tune the models for optimal performance.

The RandomForestClassifier exhibited superior performance, with an accuracy that underscored the model's robustness in handling the dataset's inherent complexities. The GradientBoostingClassifier also displayed commendable predictive capabilities, although it required careful tuning of learning rates to balance bias and variance effectively.

In conclusion, implementing machine learning models provided a foundational understanding of the stress indicators within the dataset. The RandomForest and GradientBoosting classifiers, in particular, proved to be powerful tools in stress detection, laying the groundwork for further exploration into deep learning models, which will be discussed in the subsequent section.

5.3 Implementation of Deep Learning Models

The exploration into stress detection within a healthcare environment progresses with the implementation of deep learning models. This phase aims to harness the sophisticated computational power of deep learning to uncover nuanced patterns in the physiological data that might elude traditional machine learning models.

Deep learning models necessitate a robust computational framework due to their complexity and the extensive data processing involved. For this study, TensorFlow and Keras were chosen for their comprehensive libraries and ease of use. These platforms offer a wide range of functionalities for building, training, and evaluating deep learning models.

The primary focus was on implementing an Artificial Neural Network (ANN). This model is particularly well-suited for processing and learning from the high-dimensional data present in our dataset. The architecture of the ANN was designed with multiple layers, including input, hidden, and output layers. The input layer was configured to accept the preprocessed feature set. The hidden layers, integral to the model's learning capability, comprised multiple neurons with ReLU (Rectified Linear Unit) activation functions to introduce non-linearity into the model. The output layer used a softmax function, ideal for multi-class classification, to predict the stress level categories.

Data preparation is a crucial step in deep learning. For the ANN model, the dataset underwent normalisation to ensure the input features had equal weightage. This step is critical in preventing biases towards certain features during the training process. The dataset was then partitioned into training and testing sets, with 70% of the data allocated for training and the remaining 30% for testing, ensuring a comprehensive evaluation of the model's performance.

The ANN model was trained on the prepared dataset, employing backpropagation and gradient descent algorithms to minimise the error between predicted and actual outputs. The model's hyperparameters, such as the learning rate, number of epochs, and batch size, were fine-tuned to optimise the training process. Regularisation techniques, such as dropout, were also applied to prevent overfitting, ensuring that the model generalised well to new, unseen data.

Post-training, the ANN model's performance was evaluated using various metrics. Accuracy was a primary metric, providing an overall effectiveness measure of the model in classifying stress levels. A confusion matrix was also generated to visualise the model's performance across different stress categories, offering insights into any biases or weaknesses in its predictive capabilities. Precision, recall, and F1-score were also calculated for a more nuanced understanding of the model's performance, especially in terms of its sensitivity and specificity in stress detection.

Implementing deep learning models in stress detection poses particular challenges. The complexity of neural networks demands significant computational resources and careful tuning of parameters. Balancing the depth of the network (number of layers and neurons) with the available computational resources was crucial to avoid overfitting and ensure timely training of the model. Additionally, the interpretability of deep learning models can be limited compared to traditional machine learning models, making it challenging to understand the exact features driving the predictions.

The deep learning phase of this study, mainly through the implementation of the ANN model, provided valuable insights into the potential of advanced computational techniques in stress detection. The model's ability to process complex patterns in physiological data underscored the viability of using deep learning for real-time stress monitoring in healthcare settings. The results from this phase set the stage for further exploration into hybrid ensemble approaches, combining the strengths of both machine learning and deep learning models for enhanced predictive performance.

5.4 Implementation of Ensemble Approach

The culmination of the stress detection project in the healthcare environment involves the strategic implementation of a hybrid ensemble approach. This method amalgamates the strengths of both machine learning and deep learning models, aiming to enhance the accuracy and robustness of stress level predictions.

The ensemble model was crafted by integrating the RandomForestClassifier, XGBoostClassifier, and the MLPClassifier (Artificial Neural Network). The rationale behind selecting these specific models lies in their diverse strengths and weaknesses, which, when combined, offer a comprehensive approach to stress detection.

A StackingClassifier, part of the sklearn.ensemble module was employed to create the hybrid ensemble model. This classifier functioned by using predictions from the constituent models (Random Forest, XGBoost, and ANN) as input to a final estimator, in this case, Logistic Regression. The use of Logistic Regression as the final estimator was pivotal in effectively combining the predictions from the base models.

The hybrid ensemble model underwent rigorous training on the dataset. The process involved adjusting the weights of the individual models within the ensemble to optimise their combined predictive power. The model was then tested on the reserved portion of the dataset, ensuring an unbiased evaluation of its performance.

Key metrics such as accuracy, precision, recall, and F1-score were used to assess the hybrid ensemble model's effectiveness. The ROC-AUC (Receiver Operating Characteristic - Area Under Curve) was also calculated to evaluate the model's capability to distinguish between different stress levels. A confusion matrix was generated, providing a detailed view of the model's performance across various stress categories.

The implementation of the hybrid ensemble model presented unique challenges, primarily in optimally balancing the contribution of each base model. It was crucial to ensure that the ensemble model was independent of any single model, which could lead to prediction biases. Trial-and-error methods and cross-validation techniques were employed to find the best combination of models and their respective weights.

5.5 Conclusion

The journey through the various phases of implementing machine learning and deep learning models for stress detection in healthcare settings culminated in the development of a robust hybrid ensemble model. This model synergistically combined the predictive capabilities of Random Forest, XGBoost, and ANN, resulting in a system that not only accurately predicts stress levels but also provides insights into the multifaceted nature of stress in a high-pressure environment.

Key Takeaways:

- The dataset, rich in physiological and contextual data, provided a solid foundation for the development of predictive models.
- Machine learning models offered initial insights, with Random Forest and XGBoost demonstrating solid predictive capabilities.
- The implementation of ANN highlighted the potential of deep learning in uncovering complex patterns in physiological data.
- The hybrid ensemble approach proved to be a significant advancement, leveraging the strengths of individual models to achieve higher accuracy and reliability in stress detection.

This study paves the way for further research into the integration of more sophisticated models and techniques. Continuous improvements in model architecture, feature engineering, and the incorporation of additional contextual data can enhance the predictive accuracy. Additionally, real-world deployment and continuous feedback will be instrumental in refining the models for practical applications in healthcare settings.

The implementation of this project stands as a testament to the potential of advanced data analytics and machine learning in addressing critical challenges in healthcare. The insights gained from this study can be instrumental in developing proactive strategies for stress management among healthcare professionals, ultimately contributing to improved well-being and patient care.

CHAPTER 6: RESULTS AND DISCUSSION

This chapter is dedicated to a comprehensive analysis of the results of deploying various classifiers on the dataset concerning orientation-based activity recognition. The models scrutinized range from Machine Learning algorithms, such as Random Forest and XGBoost, to Deep Learning approaches like the MLPClassifier, and finally, a Stacking Classifier representing an Ensemble approach. The data, containing features X, Y, Z, EDA, HR, and TEMP, undergoes a transformation that includes scaling and SMOTE for addressing class imbalances, leading to an equal distribution among the three classes (0, 1, 2). The classifiers' efficacy is gauged using many metrics, and their performances are visualized through confusion matrices, distribution plots, and comparative charts, offering a clear perspective on their predictive capabilities.

6.1 Performance Metrics

In assessing the performance of our classifiers, we utilize a suite of metrics, each unravelling different facets of the model's prediction accuracy and reliability.

Accuracy

A fundamental metric, accuracy, represents the proportion of true results (both true positives and true negatives) among the total number of cases examined. It is calculated as follows:

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

Precision

Precision indicates the ratio of true positives to all positives, highlighting the model's ability to return relevant instances.

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

Recall

Also known as sensitivity, recall calculates the ratio of true positives to actual positive cases, reflecting the model's capability to identify all relevant instances.

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

F1 Score

The F1 Score is the harmonic mean of precision and recall, balancing the two by considering false positives and false negatives.

$$F1\,Score = 2\times\frac{Precision \times Recall}{Precision + Recall}$$

ROC-AUC

The Receiver Operating Characteristic - Area Under Curve (ROC-AUC) metric is a performance measurement for classification problems at various threshold settings. ROC is a probability curve, and AUC represents the degree or measure of separability, depicting how much the model can distinguish between classes.

With these metrics at hand, we delve into the individual performances of each model, starting with the intricate patterns revealed by their confusion matrices and leading up to a collective visualization of their comparative strengths and weaknesses.

6.2 Confusion Matrix

The confusion matrix is a vital component of classifier evaluation, providing insight into the true vs. predicted labels. Since our classifiers predict three possible classes (0, 1, and 2), our confusion matrix is a 3x3 grid containing nine possible prediction scenarios. The axes represent the predicted and true classes, with the matrix entries indicating the counts of instances. The diagonal elements represent correct predictions, while off-diagonal elements are misclassifications. The matrix elements are defined as:

- **True Positives (TP):** Correctly predicted positive observations for the respective class.
- **True Negatives (TN):** Correctly predicted negative observations for all classes other than the respective class.
- **False Positives (FP):** Incorrectly predicted positive observations for the respective class.
- **False Negatives (FN):** Incorrectly predicted negative observations for the respective class.

For a multi-class classification scenario, these terms can be extended to TP_{ij}, TN_{ij}, FP_{ij}, and FN_{ij}, where i is the true class, and j is the predicted class.

Confusion Matrix Interpretation

A high number of TP and TN values with low FP and FN values typically indicate a well-performing model.

A high number of FP values for a class indicates that the model is prone to incorrectly predicting that class.

A high number of FN values for a class indicates that the model often misses predicting that class.

Each model's confusion matrix will be individually analyzed to extract such nuances in performance.

6.3 Label Distribution, Correlation and Feature Importance

In the initial dataset, we observed a significant class imbalance, which can adversely affect the performance of machine learning models, as they tend to be biased towards the majority class. The original label distribution was as follows:

- Class 0: 2,162,246 instances
- Class 1: 806,222 instances
- Class 2: 8,540,583 instances

Such disproportion in class distribution can lead to poor generalization of minority classes. We employed the SMOTE technique to mitigate this issue, which synthetically generates new instances in the minority classes to achieve a balanced dataset. The label distribution after applying SMOTE, ensuring that each class has an equal presence, became:

- Class 0: 2,162,246 instances
- Class 1: 2,162,246 instances
- Class 2: 2,162,246 instances

The application of SMOTE enables the models to learn more generalizable patterns, which is particularly beneficial in preventing overfitting to the majority class and under-representing the minority classes. This balanced approach is expected to improve the classification performance, evident in the improved accuracy and F1 scores in the model evaluation results.

Now, we will discuss the correlation matrix's insights and how feature interactions may influence model predictions. Additionally, the section will analyze the permutation importance graph to understand which features significantly impact model decisions.

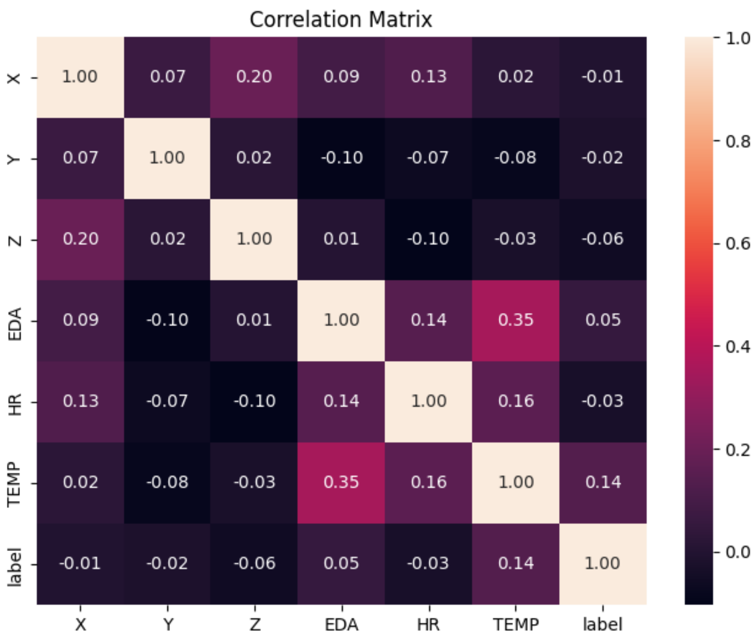


Figure 6.1: Correlation Matrix

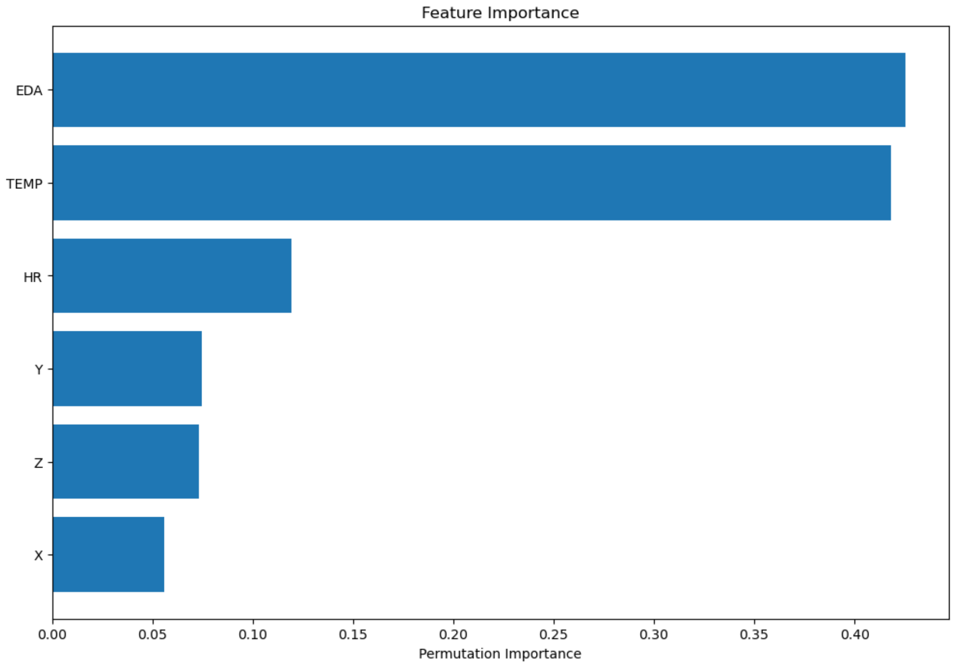
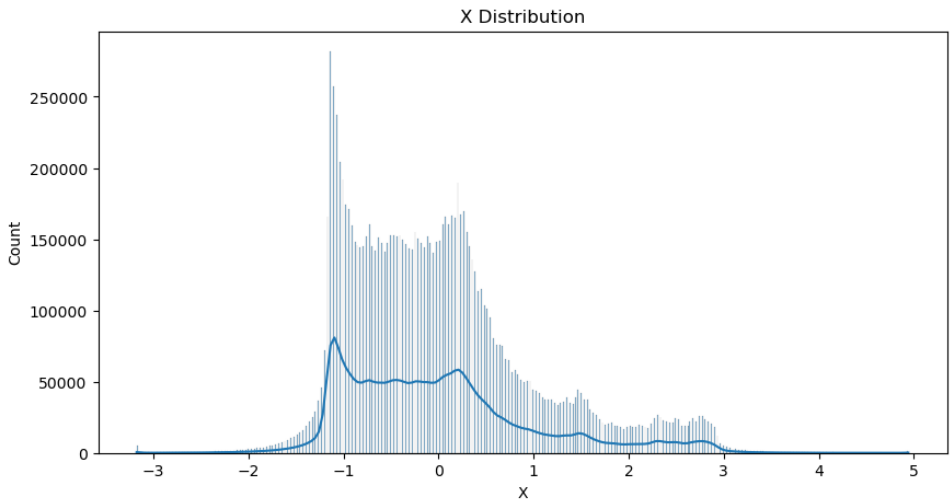
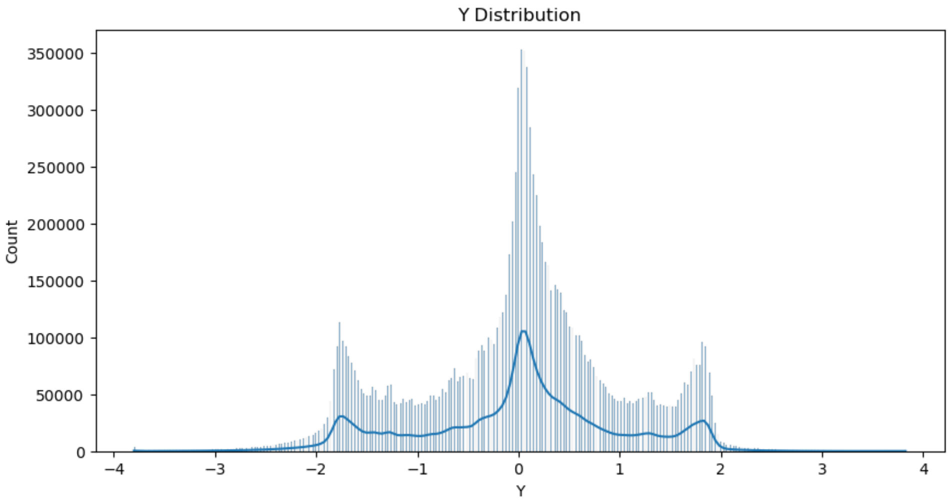


Figure 6.2: Feature Importance

6.4 Feature Analysis and Distribution

This section will focus on the distribution of features as shown in the histograms for X, Y, Z, EDA, HR, and TEMP distributions.





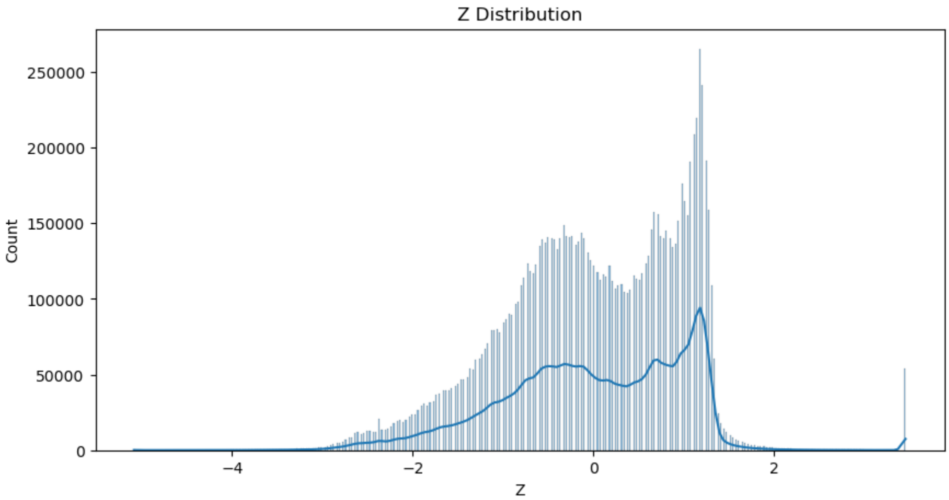
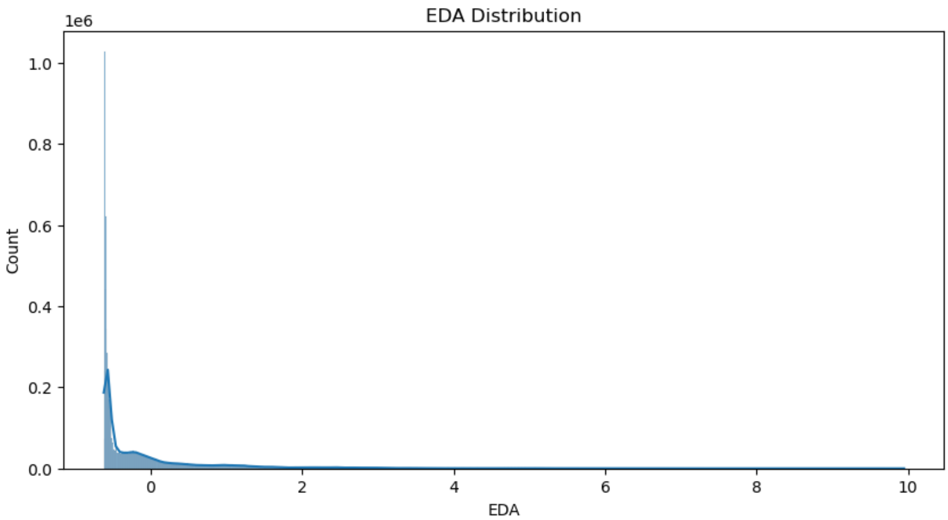


Figure 6.3: Feature Distribution (X, Y and Z)



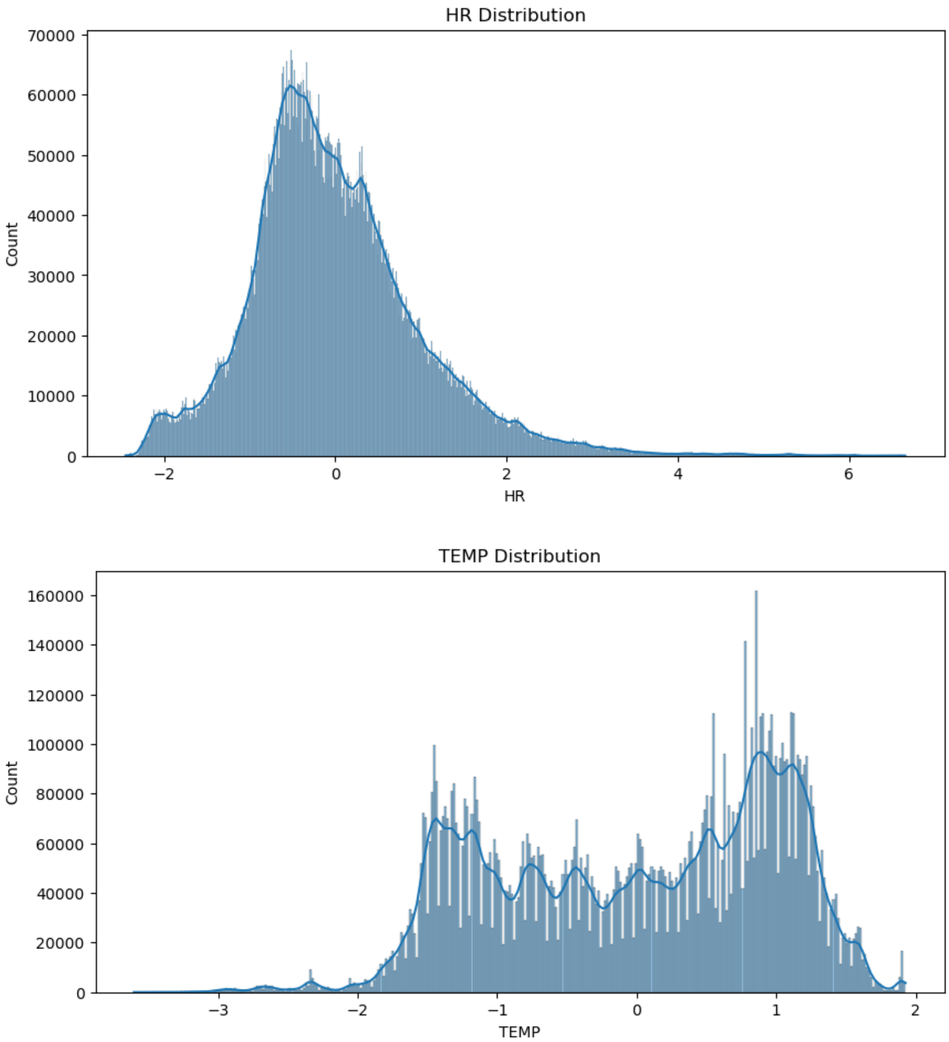


Figure 6.4: Feature Distribution (EDA, HR and TEMP)

6.5 Results of Machine Learning Models

This section presents the detailed results obtained from the Machine Learning models employed, namely Random Forest and XGBoost.

6.5.1 Random Forest

The Random Forest classifier is known for its robustness, owing to its ensemble approach in decision-making. Our Random Forest model exhibits exemplary performance with an accuracy score of 0.9934, showcasing high precision and recall across all classes.

Accuracy: 0.9934199099496306

ROC-AUC: 0.9998676251875752

Table 6.1: Classification Report of Random Forest

label	precision	recall	f1-score	support
0.0	0.99	0.99	0.99	648556
1.0	1.00	0.99	0.99	649449

2.0	0.99	0.99	0.99	648017
accuracy			0.99	1946022
macro avg	0.99	0.99	0.99	1946022
weighted avg	0.99	0.99	0.99	1946022

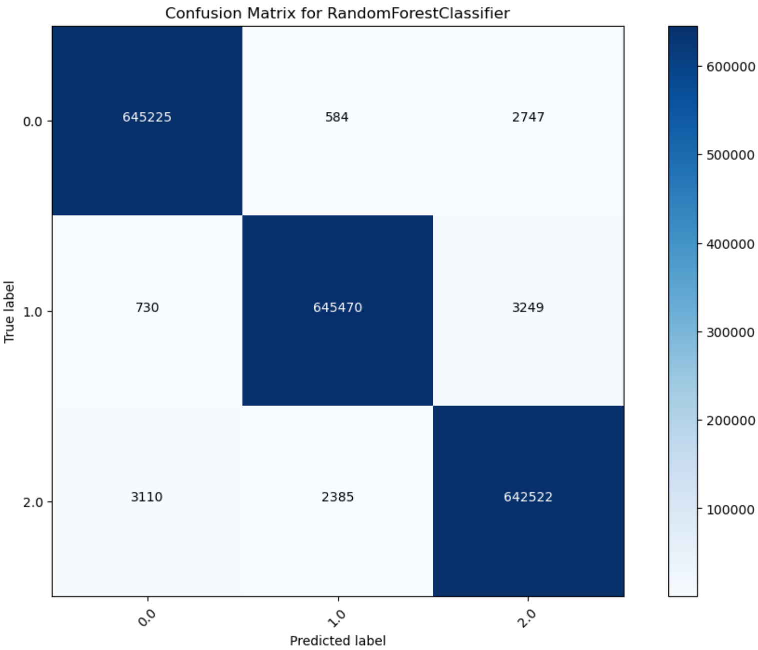


Figure 6.5: Confusion Matrix of Random Forest

6.5.2 XGBoost

XGBoost stands out for its efficiency and effectiveness on structured data. The model achieved a commendable accuracy score of 0.8833, with the classification report reflecting solid precision and recall figures.

Accuracy: 0.883350753485829

ROC-AUC: 0.9757675007454724

Table 6.2: Classification Report of XGBoost

label	precision	recall	f1-score	support
0.0	0.89	0.88	0.88	648556
1.0	0.90	0.92	0.91	649449
2.0	0.87	0.85	0.86	648017
accuracy			0.88	1946022
macro avg	0.88	0.88	0.88	1946022
weighted avg	0.88	0.88	0.88	1946022

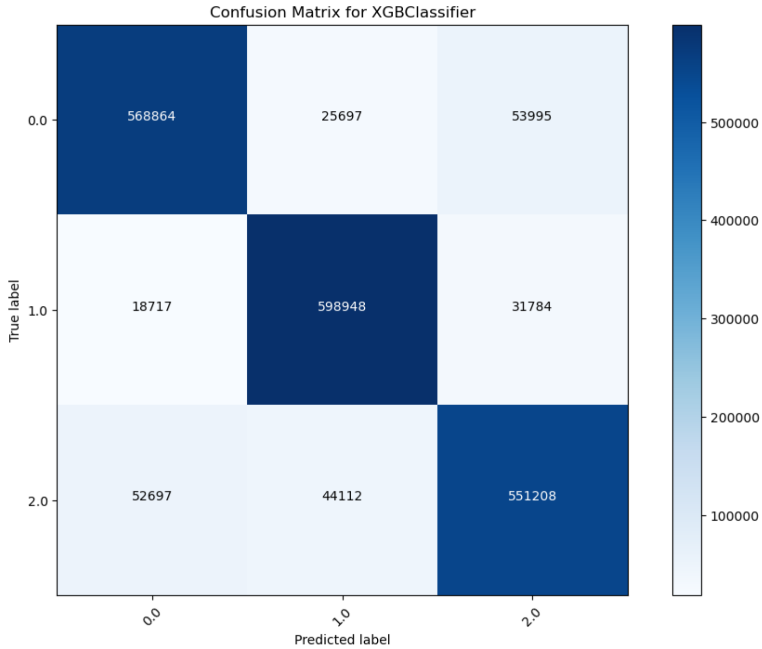


Figure 6.6: Confusion Matrix of XGBoost

6.6 Results of Deep Learning Models

Deep learning models are potent tools for pattern recognition, capable of capturing complex structures in high-dimensional data. This section presents the results obtained from the ANN model, specifically a Multilayer Perceptron (MLP) Classifier.

6.6.1 MLP Classifier

The MLP Classifier, a type of feedforward artificial neural network, has shown a promising ability to differentiate between the classes in our dataset. Despite the complexity of the task, the model achieved an accuracy of 0.7348, which, while not as high as some machine learning models, still demonstrates considerable predictive capability.

Accuracy: 0.7348169753476579

ROC-AUC: 0.8978676989050117

Table 6.3: Classification Report of MLP Classifier

label	precision	recall	f1-score	support
0.0	0.73	0.74	0.73	648556
1.0	0.78	0.77	0.77	649449
2.0	0.70	0.70	0.70	648017
accuracy			0.73	1946022
macro avg	0.73	0.73	0.73	1946022
weighted avg	0.73	0.73	0.73	1946022

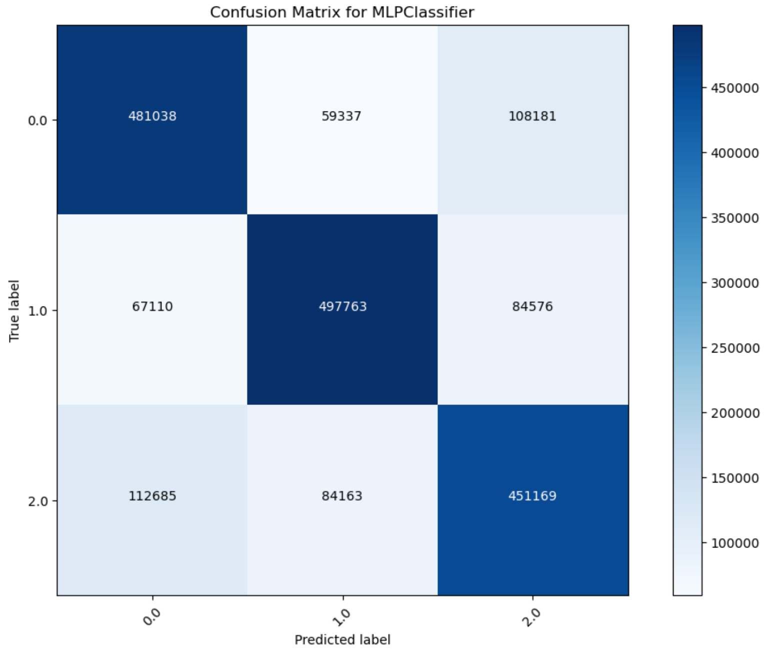


Figure 6.7: Confusion Matrix of MLP Classifier

The results underscore the trade-offs inherent in model selection: while deep learning models like MLP Classifier can capture complex relationships, they may require more data and training time to achieve the high accuracy evident in some machine learning models.

6.7 Results of Ensemble Approach

Ensemble methods, which combine several models' predictions, often improve accuracy and robustness over single-model approaches. Our implementation of a Stacking Classifier incorporates multiple learning algorithms, leading to a model that outperforms the individual classifiers in some metrics.

6.7.1 Stacking Classifier

The Stacking Classifier demonstrated exceptional performance with an accuracy score of 0.9942, indicating that the ensemble approach effectively captures the underlying patterns in the data.

Accuracy: 0.9942045876151452

ROC-AUC: 0.9998722361215356

Table 6.4: Classification Report of Stacking Classifier

label	precision	recall	f1-score	support
0.0	0.89	0.88	0.88	648556
1.0	0.90	0.92	0.91	649449
2.0	0.87	0.85	0.86	648017
accuracy			0.88	1946022
macro avg	0.88	0.88	0.88	1946022
weighted avg	0.88	0.88	0.88	1946022

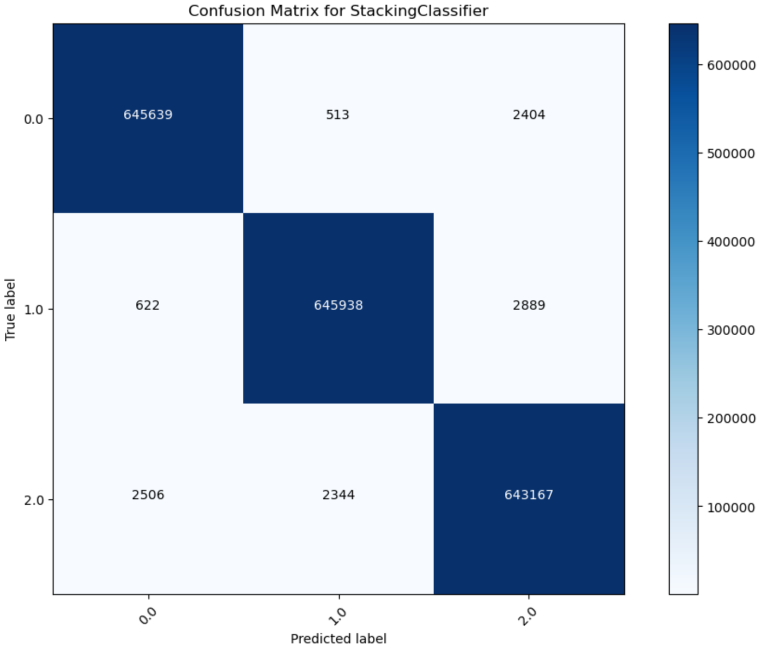


Figure 6.8: Confusion Matrix of Stacking Classifier

6.8 Comparison of Models

This section compares the performance of the various classifiers used in this study. The comparison is based on accuracy, precision, recall, F1 score, and ROC-AUC values, providing a multi-faceted view of each model's strengths and weaknesses.

Table 6.5: Comparison of Results of Different Models

Models	Random Forest	XGBoost	MLP	Stacking
Accuracy	0.9934	0.8833	0.7348	0.9942
Precision	0.99	0.88	0.73	0.99
Recall	0.99	0.88	0.73	0.99
F1-Score	0.99	0.88	0.73	0.99
ROC-AUC	0.9998	0.9758	0.8979	0.9998

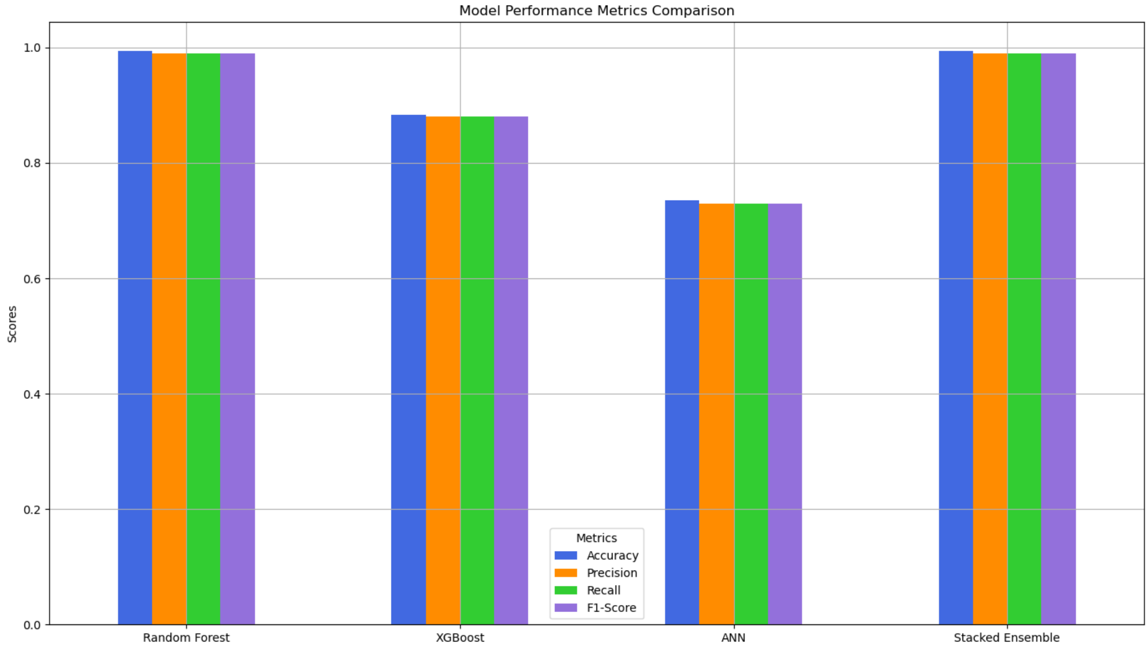


Figure 6.9: Comparison of Results of Different Models (Bar Graph)

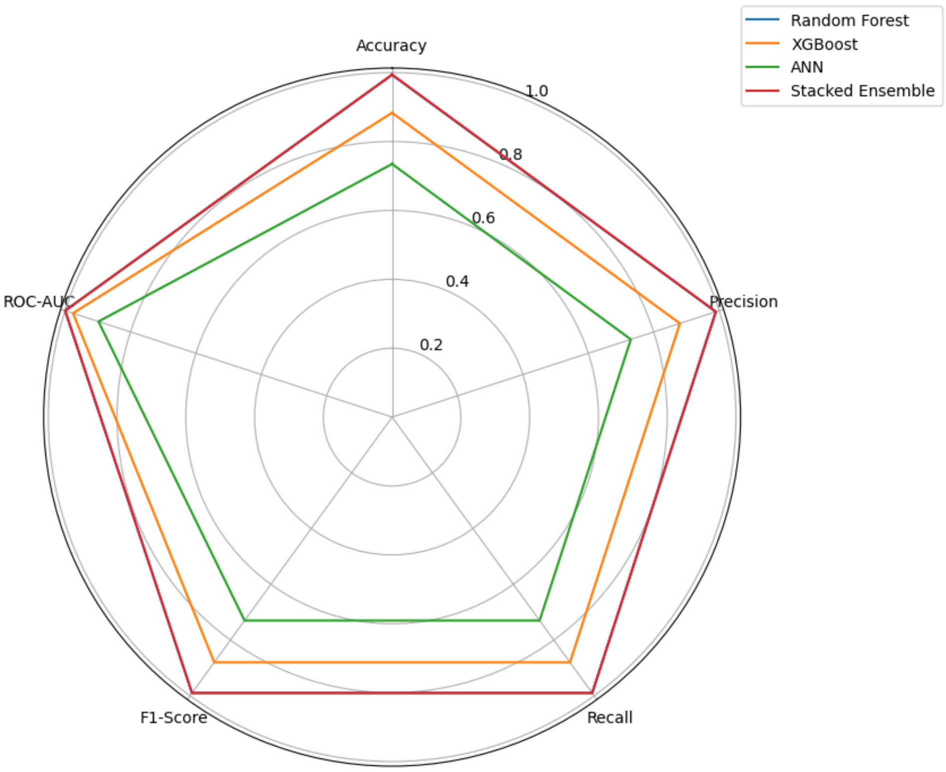


Figure 6.10: Model Performance Metrics Radar Chart

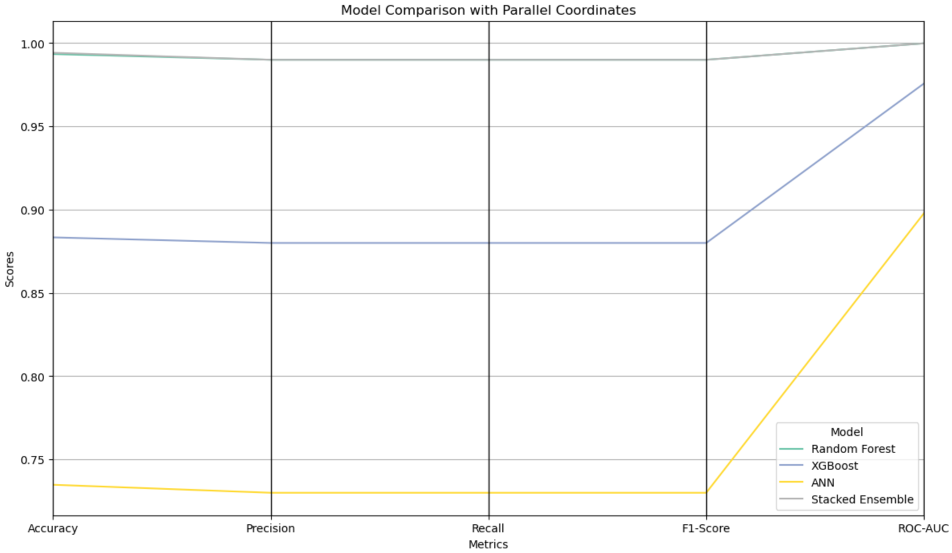


Figure 6.11: Model Performance with Parallel Coordinates

Through these comparisons, we aim to provide clear insights into which model or models may be best suited for specific types of classification tasks, especially in the context of multi-class classification challenges.

6.9 Result Discussion

In this section, we discuss the results in depth, reflecting on the implications of the performance of each model. We delve into the potential reasons behind each model's performance, considering aspects such as model complexity, data distribution, and the nature of the classification task.

This discussion will include a consideration of the following:

- The balance between precision and recall, especially in the context of our multi-class classification problem where the misclassification of different labels may have varying levels of consequence.
- The trade-offs between model complexity and performance mainly focus on whether more complex models like the MLPClassifier and StackingClassifier provide sufficient performance improvements to justify their use.
- The implications of the ROC-AUC values in the context of our specific dataset and problem and how they compare with other metrics like accuracy and F1 score.

Additionally, we will discuss the distributions of the features and their importance as provided in the images related to feature distribution and importance and how they might influence the choice of model and the interpretation of results.

Finally, we will synthesize the insights from the model comparisons to provide recommendations on model selection and deployment based on the specific needs and constraints of the stress detection task.

6.10 Conclusion

In summary, this chapter has thoroughly evaluated the classifiers utilized for the task at hand. The Random Forest and Stacking Classifier models have shown exceptional performance across all metrics, with the XGBoost and MLPClassifier following closely. The high ROC-AUC scores across the models indicate excellent separability between the classes, with the Stacking Classifier slightly outperforming others, suggesting a robust model fit.

Furthermore, the feature distributions and the 3D scatter plot offer valuable insights into the dataset's structure and the feature space within which the models operate. The importance of EDA, TEMP, and HR as significant features is particularly noteworthy, as illustrated in the feature importance graph. These features' distributions highlight the variance and skewness in the dataset, which the models have managed to capture effectively.

The model comparison charts, including the radar chart, grouped bar chart, and parallel coordinates, have provided an overview of how each model fares against the others in a clear, visual format. This comparative analysis helps understand the trade-offs in model selection, such as complexity versus interpretability and performance versus computational efficiency.

As we conclude this chapter, we reflect on the nuanced understanding these models offer for the stress detection task. The comprehensive analysis of the models and detailed feature evaluation and visualization interpretation provide a robust foundation for selecting and deploying an effective classifier for stress detection in a real-world scenario.

CHAPTER 7: CONCLUSION AND FUTURE SCOPE

7.1 Conclusion

Given the detailed analysis and discussions in the preceding chapters, our research into stress detection using wearable technology, particularly in nursing during the COVID-19 pandemic, marks a significant contribution to healthcare and wearable technology. This study aimed to leverage advanced machine learning and deep learning techniques to accurately predict stress levels from physiological data captured by wearable sensors, addressing the critical need for continuous stress monitoring in high-stress professions.

The dataset, encompassing electrodermal activity, heart rate, and temperature from nurses working in a hospital setting during the pandemic, provided a rich basis for our analyses. Through rigorous preprocessing, feature extraction, and the application of models like Random Forest, XGBoost, MLP Classifier, and a Stacking Classifier, we achieved notable accuracy, precision, recall, and F1-scores, with the Stacking Classifier model showing exceptional performance.

Our findings underscore the potential of wearable technology in real-time health monitoring and stress management, offering a non-invasive, continuous, and objective method for stress assessment. The research highlights the importance of personalized healthcare interventions and technology in enhancing individuals' well-being in critical professions.

Moreover, this study contributes to the growing body of knowledge on stress biomarkers and their correlation with various physiological parameters, paving the way for future innovations in wearable health technologies. It also reflects on the challenges and opportunities in deploying such technologies in real-world settings, emphasizing the need for scalable, robust, and user-friendly solutions.

In conclusion, this research advances our understanding of stress detection using wearable sensors and demonstrates the practical applications of machine learning and deep learning models in addressing pressing healthcare challenges. The successful prediction of stress levels among nurses during the pandemic showcases the potential of our approach to contribute significantly to occupational health, particularly in professions characterized by high-stress levels.

7.2 Future Scope

The future scope of stress detection and prediction utilizing wearable technology presents a promising frontier in personal and professional well-being. This research has unveiled the potential for significantly integrating more nuanced physiological markers, such as cortisol levels or brain activity patterns, to enhance stress detection models' accuracy. The advent of personalized stress management programs, designed to cater to individual physiological responses, heralds a new era in preventive healthcare, promising to extend its benefits beyond high-stress professions to individuals in various stressful environments, including students and those in challenging domestic situations.

The scalability of these technologies is poised to contribute to a larger, anonymized dataset that could provide invaluable insights into global stress patterns, offering a window into societal stress trends and triggers. The future also holds promise for integrating AI-driven predictive models with real-time intervention mechanisms. These could range from recommending immediate stress-relief exercises to alerting healthcare providers about individuals at risk, paving the way for more sophisticated wearable devices capable of monitoring a broader array of physiological signals and enhancing the depth and reliability of stress prediction models.

Furthermore, the advancement of wearable technology, characterized by integrating biosensors for real-time monitoring of physiological stress indicators such as heart rate variability (HRV), marks a significant leap forward. Innovations such as smart glasses and sleep trackers that analyze brainwaves underscore the potential of wearable technology to revolutionize stress management and health monitoring. The personalization of stress management solutions, powered by AI, is expected to offer highly personalized insights and recommendations, emphasizing a proactive and preventive approach to healthcare.

The exploration into mental wellness facilitated by wearable technology, which can monitor stress levels through various means and offer personalized mindfulness exercises, underscores the importance of ethical considerations, particularly regarding data privacy and security. The transition towards more inclusive and equitable healthcare solutions, facilitated by affordable and accessible wearable devices, presents an opportunity to bridge the digital health divide, offering hope to underserved communities and remote areas.

However, advancing wearable technology also brings challenges, including addressing algorithm bias and ensuring the inclusivity of health solutions. The development and implementation of these technologies must prioritize human-centred design and ethical considerations to avoid exacerbating existing inequalities.

In conclusion, the journey forward in stress detection using wearable technology is multifaceted, encompassing the advancement of technical capabilities and the creation of a holistic ecosystem that supports mental and physical health. Through continued interdisciplinary research, development, and ethical consideration, wearable technology is poised to become a cornerstone of modern healthcare strategies. Offering proactive, personalized, and preventive care solutions, these innovations promise a future where technology empowers us to manage stress more effectively and lead healthier, happier lives. The commitment to ethical, privacy, and security standards will be crucial in realizing this vision, requiring collaboration across technological, legal, and ethical domains to ensure a future where wearable technology enhances well-being in an inclusive, responsible, and impactful manner.

REFERENCES

1. Iqbal T, Simpkin AJ, Roshan D, Glynn N, Killilea J, Walsh J, Molloy G, Ganly S, Ryman H, Coen E, et al. Stress Monitoring Using Wearable Sensors: A Pilot Study and Stress-Predict Dataset. Sensors. 2022; 22(21):8135. https://doi.org/10.3390/s22218135
2. Nishtha Phutela, Devanjali Relan, Goldie Gabrani, Ponnurangam Kumaraguru, Mesay Samuel, "Stress Classification Using Brain Signals Based on LSTM Network", Computational Intelligence and Neuroscience, vol. 2022, Article ID 7607592, 13 pages, 2022. https://doi.org/10.1155/2022/7607592
3. Vos, G., Trinh, K., Sarnyai, Z., & Rahimi Azghadi, M. (2023). Ensemble machine learning model trained on a new synthesized dataset generalizes well for stress prediction using wearable devices. Journal of Biomedical Informatics, 148, 104556. https://doi.org/10.1016/j.jbi.2023.104556
4. Pratima Saravanan, Nisar, T., Zhang, Q., Masud, F., & Farzan Sasangohar. (2023). Occupational stress and burnout among intensive care unit nurses during the pandemic: A prospective longitudinal study of nurses in COVID and non-COVID units. Frontiers in Psychiatry, 14. https://doi.org/10.3389/fpsyt.2023.1129268
5. Talaat, F. M., & Rana Mohamed El-Balka. (2023). Stress monitoring using wearable sensors: IoT techniques in medical field. Neural Computing and Applications, 35(25), 18571–18584. https://doi.org/10.1007/s00521-023-08681-z
6. Purnendu Shekhar Pandey. (2017). Machine Learning and IoT for prediction and detection of stress. https://doi.org/10.1109/iccsa.2017.8000018
7. Sreedevi Uday, C. Jyotsna, & J. Amudha. (2018). Detection of Stress using Wearable Sensors in IoT Platform. https://doi.org/10.1109/icicct.2018.8473010
8. Raj, J. V., & Sarath T. (2019). An IoT based Real-Time Stress Detection System for Fire-Fighters. https://doi.org/10.1109/iccs45141.2019.9065866
9. Laavanya Rachakonda, Mohanty, S. P., Kougianos, E., Karunakaran, K., & Madhavi Ganapathiraju. (2018). Smart-Pillow: An IoT Based Device for Stress Detection Considering Sleeping Habits. https://doi.org/10.1109/ises.2018.00043
10. Setiawan, R., Budiman, F., & Wahyu Irfan Basori. (2019). Stress Diagnostic System and Digital Medical Record Based on Internet of Things. https://doi.org/10.1109/isitia.2019.8937273
11. Kumar, A., Sharma, K., & Sharma, A. (2021). Hierarchical deep neural network for mental stress state detection using IoT based biomarkers. Pattern Recognition Letters, 145, 81–87. https://doi.org/10.1016/j.patrec.2021.01.030
12. Shruti Gedam, & Paul, S. (2021). A Review on Mental Stress Detection Using Wearable Sensors and Machine Learning Techniques. IEEE Access, 9, 84045–84066. https://doi.org/10.1109/access.2021.3085502
13. Chen, J., Maysam Abbod, & Shieh, J.-S. (2021). Pain and Stress Detection Using Wearable Sensors and Devices—A Review. Sensors, 21(4), 1030–1030. https://doi.org/10.3390/s21041030

14. Shruti Gedam, & Paul, S. (2020). Automatic Stress Detection Using Wearable Sensors and Machine Learning: A Review. https://doi.org/10.1109/icccnt49239.2020.9225692
15. Tazarv, A., Labbaf, S., Reich, S. M., Dutt, N. D., Rahmani, A. M., & Levorato, M. (2021). Personalized Stress Monitoring using Wearable Sensors in Everyday Settings. 2021 43rd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC). https://doi.org/10.1109/embc46164.2021.9630224
16. Gil-Martin, M., San-Segundo, R., Mateos, A., & Ferreiros-Lopez, J. (2022). Human Stress Detection With Wearable Sensors Using Convolutional Neural Networks. IEEE Aerospace and Electronic Systems Magazine, 37(1), 60–70. https://doi.org/10.1109/maes.2021.3115198
17. Sethi, Kriti & Ramya, T. & Singh, Hanut & Dutta, Rishik. (2019). Stress detection and relief using wearable physiological sensors. TELKOMNIKA (Telecommunication Computing Electronics and Control). 17. 1139. https://doi.org/10.12928/telkomnika.v17i3.9719
18. Amr Gaballah, Tiwari, A., Narayanan, S., & Falk, T. H. (2021). Context-Aware Speech Stress Detection in Hospital Workers Using Bi-LSTM Classifiers. https://doi.org/10.1109/icassp39728.2021.9414666
19. Sakib, M., & Syeda Shanaz Pervez. (2023). Automated Stress Level Detection for Hospital Nurses: A Single Triaxial Wearable Accelerometer Sensor System Approach. https://doi.org/10.1109/cce60043.2023.10332832
20. Morales, A., Barbosa, M., Morás, L., Silvio César Cazella, Sgobbi, L. F., Sene, I., & Marques, G. (2022). Occupational Stress Monitoring Using Biomarkers and Smartwatches: A Systematic Review. Sensors, 22(17), 6633–6633. https://doi.org/10.3390/s22176633
21. Hosseini, S., Katragadda, S., Ravi Teja Bhupatiraju, Ashkar, Z., Borst, C. W., Cochran, K., & Raju Gottumukkala. (2022). A multimodal sensor dataset for continuous stress detection of nurses in a hospital. Scientific Data, 9(1). https://doi.org/10.1038/s41597-022-01361-y
22. Kinnaird L. (2015). Nurses: The Backbone of Health Services. The Florida Nurse, 63(4). Retrieved from https://pubmed.ncbi.nlm.nih.gov/26817189/
23. Mariani, M. M., Machado, I., Vittoria Magrelli, & Dwivedi, Y. K. (2023). Artificial intelligence in innovation research: A systematic review, conceptual framework, and future research directions. Technovation, 122, 102623–102623. https://doi.org/10.1016/j.technovation.2022.102623
24. Heckman, W. (2019). How the Fight or Flight Response Works - The American Institute of Stress. Retrieved February 1, 2024, from The American Institute of Stress website: https://www.stress.org/how-the-fight-or-flight-response-works
25. Yartsev, A. (2023, December 18). Physiological response to changes in posture. Retrieved February 1, 2024, from Deranged Physiology website: https://derangedphysiology.com/main/cicm-primary-exam/required-

reading/cardiovascular-system/Chapter%20501/physiological-response-changes-posture
26. Fostering a Culture of Belonging in the Hybrid Workplace. (2021, August 3). Retrieved February 1, 2024, from Harvard Business Review website: https://hbr.org/2021/08/fostering-a-culture-of-belonging-in-the-hybrid-workplace
27. Liljestrand, R., & Martin, S. (2021). Stress and Resilience Among Healthcare Workers During the COVID-19 Pandemic. Rehabilitation Nursing Journal, 46(6), 300–304. https://doi.org/10.1097/rnj.0000000000000344
28. Xu, C., Song, Y., Sempionatto, J. R., Solomon, S. A., Yu, Y., Nyein, H. Y., Tay, R. Y., Li, J., Heng, W., Min, J., Lao, A., Hsiai, T. K., Sumner, J. A., & Gao, W. (2024). A physicochemical-sensing electronic skin for stress response monitoring. Nature Electronics, 1-12. https://doi.org/10.1038/s41928-023-01116-6
29. Dataquest. (2023, March 31). Tutorial: Introduction to Deep Learning. Retrieved February 1, 2024, from Dataquest website: https://www.dataquest.io/blog/tutorial-introduction-to-deep-learning/
30. Rab, S. L., & Admon, R. (2020). Parsing inter- and intra-individual variability in key nervous system mechanisms of stress responsivity and across functional domains. Neuroscience & Biobehavioral Reviews, 120, 550-564. https://doi.org/10.1016/j.neubiorev.2020.09.007
31. Crina Grosan, & Abraham, A. (2011). Rule-Based Expert Systems. Intelligent Systems Reference Library, 149–185. https://doi.org/10.1007/978-3-642-21004-4_7
32. Brownlee, J. (2021, May 11). A Gentle Introduction to Multiple-Model Machine Learning - MachineLearningMastery.com. Retrieved February 1, 2024, from MachineLearningMastery.com website: https://machinelearningmastery.com/multiple-model-machine-learning/
33. Prasad, K., McLoughlin, C., Stillman, M., Poplau, S., Goelz, E., Taylor, S., … Sinsky, C. A. (2021). Prevalence and correlates of stress and burnout among U.S. healthcare workers during the COVID-19 pandemic: A national cross-sectional survey study. EClinicalMedicine, 35, 100879–100879. https://doi.org/10.1016/j.eclinm.2021.100879
34. Armstrong, T., Wells, J., Boisvert, D. L., Lewis, R., Cooke, E. M., Woeckener, M., & Kavish, N. (2019). Skin conductance, heart rate and aggressive behavior type. Biological Psychology, 141, 44–51. https://doi.org/10.1016/j.biopsycho.2018.12.012
35. Ramaekers, D., Ector, H., Aubert, A. E., Rubens, A., & Van. (1998). Heart rate variability and heart rate in healthy volunteers. Is the female autonomic nervous system cardioprotective? European Heart Journal, 19(9), 1334–1341. https://doi.org/10.1053/euhj.1998.1084
36. Nelli, F. (2023). Python Data Analytics. SpringerLink. https://doi.org/10.1007-978-1-4842-9532-8

37. Vaibhav Jayaswal. (2020, September 7). Performance Metrics: Accuracy - Towards Data Science. Retrieved February 1, 2024, from Medium website: https://towardsdatascience.com/performance-metrics-accuracy-118f728d328f
38. Koehrsen, W. (2018, June 22). A Feature Selection Tool for Machine Learning in Python. Retrieved February 1, 2024, from Medium website: https://towardsdatascience.com/a-feature-selection-tool-for-machine-learning-in-python-b64dd23710f0
39. Evaluation Metrics in Machine Learning. (2020, September). Retrieved February 1, 2024, from GeeksforGeeks website: https://www.geeksforgeeks.org/metrics-for-machine-learning-model/
40. Hasan, B. (2023, November 16). What Is Random Forest In Machine Learning | Robots.net. Retrieved February 1, 2024, from Robots.net website: https://robots.net/fintech/what-is-random-forest-in-machine-learning/
41. The Click Reader. (2021, October 19). Random Forest Regression Explained with Implementation in Python. Retrieved February 1, 2024, from Medium website: https://medium.com/@theclickreader/random-forest-regression-explained-with-implementation-in-python-3dad88caf165
42. Oguz, B. U., Shinohara, R. T., Yushkevich, P. A., & Oguz, I. (2017). Gradient Boosted Trees for Corrective Learning. Lecture Notes in Computer Science, 203–211. https://doi.org/10.1007/978-3-319-67389-9_24
43. Carlos Rodriguez (he/him. (2023, May 20). The Notorious XGBoost - Towards Data Science. Retrieved February 6, 2024, from Medium website: https://towardsdatascience.com/the-notorious-xgboost-c7f7adc4c183
44. Do you know what deep learning is? (2022). Retrieved February 1, 2024, from Oracle.com website: https://www.oracle.com/artificial-intelligence/machine-learning/what-is-deep-learning/
45. Basheer, I. A., & M Hajmeer. (2000). Artificial neural networks: fundamentals, computing, design, and application. Journal of Microbiological Methods, 43(1), 3–31. https://doi.org/10.1016/s0167-7012(00)00201-3
46. Wikipedia Contributors. (2024, January 15). Ensemble learning. Retrieved February 1, 2024, from Wikipedia website: https://en.wikipedia.org/wiki/Ensemble_learning
47. Vadim Smolyakov. (2017, August 22). Ensemble Learning to Improve Machine Learning Results. Retrieved February 1, 2024, from Medium website: https://medium.com/cube-dev/ensemble-learning-d1dcd548e936
48. Raschka, S. (2015). StackingClassifier: Simple stacking - mlxtend. Retrieved February 1, 2024, from Github.io website: https://rasbt.github.io/mlxtend/user_guide/classifier/StackingClassifier/
49. Chawla, N. V., Bowyer, K. W., Hall, L. O., & W. Philip Kegelmeyer. (2002). SMOTE: Synthetic Minority Over-sampling Technique. Journal of Artificial Intelligence Research, 16, 321–357. https://doi.org/10.1613/jair.953
50. Pratheesh Shivaprasad. (2020, October 19). Understanding Confusion Matrix, Precision-Recall, and F1-Score. Retrieved February 1, 2024, from Medium

website: https://towardsdatascience.com/understanding-confusion-matrix-precision-recall-and-f1-score-8061c9270011
51. Gad, A. (2018, September 11). Deduce the Number of Layers and Neurons for ANN. Retrieved February 1, 2024, from Datacamp.com website: https://www.datacamp.com/tutorial/layers-neurons-artificial-neural-networks
52. Mohammed, A., & Kora, R. (2023). A comprehensive review on ensemble deep learning: Opportunities and challenges. Journal of King Saud University - Computer and Information Sciences, 35(2), 757–774. https://doi.org/10.1016/j.jksuci.2023.01.014
53. Great Learning Team. (2020, July 21). Understanding ROC (Receiver Operating Characteristic) Curve | What is ROC? Retrieved February 1, 2024, from Great Learning Blog: Free Resources what Matters to shape your Career! website: https://www.mygreatlearning.com/blog/roc-curve/
54. Model. (2020, February 19). DataSklr. Retrieved February 1, 2024, from DataSklr website: https://www.datasklr.com/select-classification-methods/model-selection