STRESS CLASSIFICATION FROM WEARABLE MULTIMODAL PHYSIOLOGICAL SIGNALS: DEEP LEARNING APPROACH

BY

Aninda Halder

ID: 0242310005103017

This Report Presented in Partial Fulfillment of the Requirements for
The Degree of Master of Science in Computer Science and Engineering

Supervised By

Md. Sadekur Rahman

Assistant Professor

Department of CSE

Daffodil International University

DAFFODIL INTERNATIONAL UNIVERSITY

DHAKA, BANGLADESH

July 2024

APPROVAL

This Project titled "**Stress Classification From Wearable Multimodal Physiological Signals: Deep Learning Approach**", submitted by **Aninda Halder** to the Department of Computer Science and Engineering, Daffodil International University, has been accepted as satisfactory for the partial fulfillment of the requirements for the degree of M.Sc. in Computer Science and Engineering (MSc.) and approved as to its style and contents. The presentation has been held on 06/07/2024.

BOARD OF EXAMINERS

Chairman

Dr. Sheak Rashed Haider Noori, PhD
Professor and Head
Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University

Internal Examiner

Dr. Md. Zahid Hasan, PhD
Associate Professor
Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University

Internal Examiner

Dr. Arif Mahmud, PhD
Associate Professor
Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University

External Examiner

Dr. Md. Zulfiker Mahmud
Professor
Department of Computer Science and Engineering
Jagannath University

DECLARATION

We hereby declare that this thesis has been done by us under the supervision of **Md. Sadekur Rahman, Assistant Professor, Department of CSE,** Daffodil International University. We also declare that neither this thesis nor any part of this thesis has been submitted elsewhere for award of any degree or diploma.

Supervised by:

Md. Sadekur Rahman
Assistant Professor
Department of CSE
Daffodil International University

Submitted by:

Aninda Halder
ID: 0242310005103017
Department of CSE
Daffodil International University

ACKNOWLEDGEMENT

First, we express our heartiest thanks and gratefulness to almighty God for His divine blessing makes us possible to complete the final thesis successfully.

We are grateful and wish our profound indebtedness to **Md. Sadekur Rahman, Assistant Professor**, Department of CSE Daffodil International University, Dhaka. Deep Knowledge & keen interest of our supervisor in the field of "Data Mining" to carry out this thesis. His endless patience, scholarly guidance, continual encouragement, constant and energetic supervision, constructive criticism, valuable advice, reading many inferior drafts and correcting them at all stages have made it possible to complete this thesis.

We would like to express our heartiest gratitude to **Dr. Sheak Rashed Haider Noori, Professor and Head, Department of CSE**, for his kind help to finish our thesis and to other faculty members and the staff of CSE department of Daffodil International University.

We would like to thank our entire course mate in Daffodil International University, who took part in this discussion while completing the course work.

Finally, we must acknowledge with due respect the constant support and passion of our parents.

ABSTRACT

Stress is a common problem with serious consequences for health and well-being, especially among healthcare workers such as nurses. Chronic stress in this population can lead to burnout, affecting both their personal well-being and patient care. Traditional stress assessment methods are limited in their ability to provide continuous, objective monitoring, highlighting the need for innovative solutions. Wearable technology, such as the `Empatica E4` smartwatch, offers a promising avenue for real-time stress monitoring through the measurement of physiological signals. In our study, we developed a stress classification model utilizing data collected from the `Empatica E4` smartwatch worn by nurses in a hospital setting. The model uses physiological signals to effectively classify stress levels using Long Short-Term Memory (LSTM) neural networks. The study begins by exploring the physiological and psychological aspects of stress, emphasizing the challenges faced by nurses in high-stress environments and the potential of wearable technology for stress monitoring. Utilizing a publicly available dataset consisting of physiological signals, including heart rate, electrodermal activity (EDA), skin temperature, and the physical activity of nurses, and preprocessing it for analysis. The results demonstrate the model's effectiveness in both binary and multiclass classification tasks, with notable performance differences observed across stress categories. In multiclass classification, the model exhibits high precision and recall for "no stress" (precision: 0.88, recall: 0.90) and "high stress" (precision: 0.94, recall: 0.94) categories but shows reduced performance for "low stress" (precision: 0.74, recall: 0.71). This gap could be related to signal ambiguity and class imbalance. In binary classification, merging "low stress" with "no stress" simplifies the task and results in consistent performance across both classes, with stress detection achieving precision of 0.91 and recall of 0.96 and "no stress" classification precision of 0.90 and recall of 0.82. This research not only advances our understanding of stress monitoring in the healthcare environment but also furthers the exploration of stress diagnosis using smartwatch data. The developed stress classification model offers potential applications in real-time stress monitoring systems, personalized feedback mechanisms, and intervention strategies to mitigate stress among healthcare professionals. These findings have broader implications for improving well-being and patient care quality in high-stress environments beyond healthcare.

Table of Contents

List Of Tables

List Of Figures

CHAPTER 1
INTRODUCTION

1.1 Introduction

Stress is common in the human body and has many different aspects. It strongly affects our health and well-being. Stress is the body's response to challenges or threats. In the context of physiology, stress triggers a series of responses to tackle challenges. And these responses boost the levels of hormones like cortisol and adrenaline, governed by the autonomic nervous system [1] [2]. This response, known as the "fight-or-flight" reaction, and the release of these hormones lead to a rise in heart rate, blood pressure, and elevated levels of energy [3]. However, when stress lasts for a long time or causes a chronic stress state, it can harm one's health. It can cause heart disease, anxiety, depression, a weak immune system, faster biological aging, etc.

In healthcare settings, stress is especially high among healthcare professionals, particularly nurses. For patient care, nurses often work long hours in high-stakes environments where they must make quick decisions and handle multiple tasks simultaneously. The kind of work they do means they deal with stress all the time. They must handle very sick patients and deal with relationships with colleagues and patients' families. Leading to chronic stress, causing emotional exhaustion, depersonalization, and reduced personal accomplishment. These not only affect the well-being of nurses but also compromise patient care, leading to increased medical errors, decreased patient satisfaction, and higher turnover rates among nursing staff [4].

Because stress affects healthcare workers so much, it's really important to keep an eye on it and find good ways to deal with it. Traditional methods of stress assessment are often self-report questionnaires and periodic psychological evaluations, which are limited due to their subjective nature [5]. In contrast, wearable technology offers a promising solution for continuous and objective monitoring of physiological signals associated with stress. Devices like the Empatica E4 smartwatch can measure various physiological signals. It has different sensors [6], Photoplethysmography (PPG) to measure Blood Volume Pulse (BVP), electrodermal activity (EDA) sensor to measure change in conductivity in skin, temperature sensor to measure skin temperature, and accelerometer (ACC) sensor to measure physical activity. These parameters provide valuable insights into an individual's stress levels and can be used to develop real-time stress monitoring systems [7].

In this research, we focus on the development of a stress classification model using data collected from the Empatica E4 smartwatch worn by nurses in a hospital environment. Developing a Deep Learning model Long Short-Term Memory (LSTM) neural networks. LSTM networks are a type of recurrent neural network (RNN) [8]. LSTM networks are specifically designed to process sequential data like time-series data, natural language processing (NLP), etc., making them ideal for capturing the temporal dynamics of physiological signals and NLP as such.

With the background and motivation for this study, this paper presents an outline for the stress research objectives and formulates the research questions guiding our investigation. We begin by discussing the physiological and psychological aspects of stress, the specific challenges faced by nurses, and the potential of wearable technology for stress monitoring. With the methodology for data collection, preprocessing, and model development, followed by an evaluation of our stress classification model and its implications for healthcare.

1.2 Motivation

The impact of stress on healthcare professionals, particularly nurses, is the motivation for this study. And the potential benefits of using wearable technology and machine learning for stress management. Nurses are very important in the healthcare system, providing critical care and support to patients. However, the demanding and often high-pressure nature of their work environment places them at significant risk of stress and burnout. Studies have shown that chronic stress among nurses can lead to a range of adverse outcomes, including decreased job satisfaction, increased absenteeism, and higher rates of turnover [4]. These issues hamper patient care as well as the well-being of nurses.

The traditional methods for assessing and managing stress need to be revised to address the needs of healthcare professionals. Self-report questionnaires and periodic evaluations are subjective and may not capture the real-time dynamics of stress. There is a need for continuous, objective monitoring of stress levels to enable timely interventions. Wearable devices like the Empatica E4 smartwatch offer a promising solution. These devices keep track of one’s body's signals all the time. They give a lot of information that can show when stress levels change. Then, they can help before the stress gets too bad.

Analyzing time-series physiological signals from wearable devices using LSTM networks presents a novel approach to stress classification. LSTM networks are designed in such a way to understand long-term dependencies in sequential data, which is also promising for capturing the complex temporal patterns in physiological signals associated with stress [8]. This research aims to use LSTM networks to develop an accurate and robust stress classification model that can be deployed in real-world healthcare settings.

1.3 Research Objective

Developing a reliable and accurate stress classification model using physiological data collected from the Empatica E4 smartwatch in a healthcare environment. To achieve this objective, the study is structured around the following specific aims:

1. **Data Collection and Preprocessing:** To collect high-quality physiological data from nurses in a hospital environment and preprocess it for analysis. Hosseini et al. [9] did the first half; I utilized their publicly available data collected from a hospital environment. The other half involves handling missing data, normalizing signals, and segmenting the data into relevant time frames based on self-reported stress levels.

2. **Model Development, Training, Evaluation and Validation:** My goal is to develop an LSTM neural network model for stress classification. I have utilized the preprocessed and labeled data, as mentioned earlier, to recognize patterns associated with different stress levels. To evaluate the performance of the model I have utilized different metrics commonly used for classification assessment. And those metrics are accuracy, precision, recall, and F1-score.

3. **Interpretation and Application:** To interpret the results and application of real-world environments. This includes exploring how the model can be integrated into stress management interventions for healthcare professionals. Potential applications include real-time stress monitoring systems, personalized feedback mechanisms, and intervention strategies to mitigate stress.

1.4 Research Question

The primary research question guiding this study is:

How accurately can an LSTM neural network model classify stress levels using physiological data collected from the Empatica E4 smartwatch in a hospital environment?

To address this primary question, the study explores several sub-questions:

1. **What are the key physiological signals and features that are most indicative of stress in nurses?**
   This sub-question aims to identify the most relevant physiological markers of stress by analyzing the data collected from the Empatica E4 smartwatch. Understanding which signals and features are most correlated with stress.

2. **How do LSTM architectures differ in their performance for binary classification tasks compared to multi-class classification tasks?**
   This sub-question evaluates the effectiveness of the LSTM model. By comparing performance metrics, we can determine the advantages and limitations of using LSTM networks for stress classification.

3. **What are the practical implications in healthcare settings?**
   This sub-question explores the broader implications of implementing the developed stress classification model in real-world healthcare environments. It considers the potential benefits, challenges, and ethical considerations of using wearable technology and machine learning for continuous stress monitoring and intervention.

By addressing these questions, this research aims to contribute to the ongoing research on stress monitoring and management using wearable technology and machine learning. Enabling the well-being of healthcare professionals and positive impact on patient care. Moreover, the methodologies and insights gained from this study can be applied to other high-stress professions and environments, broadening the impact of this research.

CHAPTER 2
BACKGROUND

2.1 Introduction

Stress monitoring and classification have become increasingly important due to the widespread recognition of stress's adverse effects on health and productivity. The development of effective stress detection systems has been driven by advances in wearable technology and machine learning. This section provides an overview of the background and related work in the field of stress classification, particularly focusing on physiological signal analysis, the use of wearable devices, and the application of machine learning and deep learning techniques. We will discuss the scope of the problem, the challenges encountered, and the contributions of previous studies that have paved the way for current research.

2.2 Related Work

2.2.1 Previous Study

In [10], authors investigate how machine learning and deep learning can detect stress using data from wearable sensors that monitor body signals like heart rate and muscle activity. They used the 'Wearable Stress and Affect Detection (WESAD)' [11] dataset, which includes various physiological signals collected during different activities like relaxation, neutral states, and stress. The study tested different algorithms and found that machine learning methods achieved accuracies of up to 81.65% for distinguishing between these states and even higher for simpler classifications. The deep learning model performed even better, reaching accuracies of 84.32% and 95.21% respectively. This suggests that deep learning shows promise in accurately identifying stress. The paper concludes by highlighting the need for personalized models since stress responses vary from person to person. Zhu et al. [12], in their study addresses the role of wearable devices, particularly smartwatches with electrodermal activity (EDA) sensors, in predicting and managing stress. It explores the effectiveness of five machine learning classifiers in distinguishing stress from non-stress states using EDA data from four databases. The Support Vector Machine (SVM) achieved the highest accuracy at 92.9% for stress prediction. Gender differences in stress classification were also noted. The research highlights the potential of multimodal approaches for enhancing mental health monitoring through wearable technology, underscoring the importance of integrating physiological data to develop effective stress management tools. Zainudin et al. [13], investigates the application of machine learning and deep learning algorithms for stress detection, emphasizing early identification to mitigate health risks. Utilizing data from the MySignals Healthcare Toolkit at the Indian Institute of Information Technology, the study collected Galvanic Skin Response (GSR) and Electrocardiogram (ECG) data from 252 participants. Among the algorithms tested, the Decision Tree (DT) classifier emerged as the most effective, achieving impressive metrics: 95% accuracy, 96% precision, recall, and F1-score. This highlights DT's robust performance in stress detection using physiological signals, underscoring its suitability for practical applications without common issues like underfitting or overfitting. Rajdeep et al. [14] in their study, investigates a wearable system's effectiveness in monitoring stress among older adults through machine learning techniques. Utilizing Electrodermal Activity (EDA) and Blood Volume Pulse (BVP) signals, correlated with salivary cortisol levels as stress biomarkers, data from 19 healthy older adults (mean age 73.15 ± 5.79 years) during the Trier Social Stress Test were analyzed. Extracting 39 statistical features from EDA and BVP signals, the study found that Logistic Regression achieved strong performance, with macro-average and micro-average F1-scores of 0.87 and 0.95, respectively, and an AUC score of 0.813. Additionally, a novel LSTM-based deep learning classifier outperformed logistic regression, showing improvements of 6.7% and 2% in macro-average and micro-average F1-scores, respectively, with an AUC score of 0.9, 11% higher than logistic regression. Ritu et al. [15] in their study, introduced a novel hybrid deep learning model combining Convolutional Neural Network (CNN) and Long Short-Term Memory (LSTM) architectures to classify stress levels based on physiological signals from wearable devices. Using the publicly available WESAD dataset from 15 participants, which includes multivariate time-series data from chest-worn and wrist-worn RespiBAN and Empatica E4 wearables, the study achieved significant results. The CNN-LSTM model demonstrated an impressive accuracy of 90.20% for stress classification, surpassing previous studies. Evaluation metrics such as confusion matrix, accuracy, precision, and F1-score were utilized to validate the model's performance, highlighting its potential in advancing stress management and healthcare systems through robust and effective stress recognition capabilities. B. Roy et al. [16] in their study introduces an innovative approach for stress detection using EEG signals, employing a hybrid deep learning model that combines Convolutional Neural Network (CNN) and Bidirectional Long Short-Term Memory (BLSTM). The study utilized the Physionet EEG dataset, focusing on mental arithmetic tasks as stress-inducing stimuli. Results demonstrate that the proposed CNN-BLSTM model achieved outstanding classification accuracy, reaching 99.20% in initial testing and 98.10% through stratified tenfold cross-validation.

2.2.2 Physiological Signals and Stress

Physiological changes have a relationship with stress; key physiological signals for stress identification are heart rate (HR), heart rate variability (HRV), respiration rate (RESPR), electrodermal activity (EDA), skin temperature (TEMP), etc. The autonomic nervous system is responsible for these changes due to stress.

1. **Heart Rate and Heart Rate Variability (HRV):** HRV is the signal that is extracted from the heartbeats with time intervals. HRV is widely used for stress identification, and research has shown that stress typically reduces HRV. Studies like those by [17] and [18] have established HRV as a reliable metric for stress detection.

2. **Electrodermal Activity (EDA):** Electrical conductivity is measured by the EDA sensor, changes in EDA of skin increase sweat gland activity driven by sympathetic nervous system arousal. Vasil et al. [19] utilized EDA to classify stress states.

3. **Skin Temperature:** Stress-induced changes in skin temperature are linked to peripheral vasoconstriction controlled by the autonomic nervous system [20]. Research by Vinkers et al. [21] demonstrated that skin temperature decreases during acute stress.

2.2.3 Wearable Devices for Stress Monitoring

Wearable technology has revolutionized the monitoring of physiological signals, providing continuous measurement capabilities. Devices like the Empatica E4 smartwatch have become integral to stressful research due to their ability to collect high-resolution data in real-time.

1. **Empatica E4 Smartwatch:** The Empatica E4 is widely used in stress research due to its capability to measure multiple physiological parameters, including heart rate, EDA, skin temperature, and movement [22].

2. **Other Wearable Devices:** Various other wearables, such as the Apple Watch, Fitbit, and Garmin devices, have also been used for stress monitoring [23].

2.2.4 Machine Learning for Stress Classification

Machine learning techniques brought a big leap in data analysis, in physiological data analysis for stress classification is even more. The ability of machine learning models to work on large datasets makes machine learning a powerful tool for stress detection and extracting meaningful patterns from the data.

1. **Traditional Machine Learning Models:** Traditional machine learning models have been widely used for stress identification, such models include Support Vector Machines (SVM), Random Forests, Decision Trees (DT), and k-Nearest Neighbors (k-NN), etc. Studies by Sharma et al. [24] used these models to classify stress based on physiological data, achieving reasonable accuracy levels.

2. **Deep Learning Models:** More recent research has focused on deep learning models. There are huge advancements in this field; models like Convolutional Neural Networks (CNNs) and Recurrent Neural networks (RNNs) are groundbreaking in the field of deep learning. The advancement of RNN shows great results with Long Short-Term Memory (LSTM) networks. Deep learning models have proven themselves in complex patterns, as well as in time-series data. For instance, studies by Phutela et al. [25] and Rani et al. [26] worked on LSTM networks and showed their effectiveness in stress classification using physiological signals, including brain signals.

3. **Hybrid Models:** Combining multiple machine learning techniques has also been explored to enhance stress classification accuracy. Research by Kang et al. [27] and Acikmese et al. [28] utilized a hybrid model integrating CNNs and LSTMs on physiological data and achieved high classification performance capturing spatial and temporal features.

2.3 Scope of the Problem

2.3.1 Stress among Healthcare Professionals

The prevalence of stress among healthcare professionals, particularly nurses, is a critical issue. Nurses’ jobs are always filled with extensive pressure due to the nature of their work, long hours, and the need to manage critical patient conditions which induce high levels of stress.

This stress can lead to burnout, adversely affecting their mental and physical health, job satisfaction, and patient care.

1. **Impact on Health:** Chronic stress in nurses can lead to various health problems, including cardiovascular diseases, depression, anxiety, and musculoskeletal disorders. Studies by Ito, et al. [29] and Mohanty [30] highlight the health risks associated with prolonged stress in healthcare settings.

2. **Impact on Job Performance:** Stress negatively impacts nurses' cognitive functions, decision-making abilities, and overall job performance. Research by Williams et al. [31] and Khamisa et al. [32] demonstrates the correlation between high stress levels and increased medical errors, reduced patient satisfaction, and higher turnover rates.

3. **Economic Impact:** The economic implications of stress-related issues in healthcare are significant, including costs associated with absenteeism, reduced productivity, and high turnover. Studies by Selamu et al. [33], Joshi et al. [34] and Wang et al. [35] discuss the financial burden of stress on healthcare systems.

2.3.2 Need for Real-Time Stress Monitoring

Traditional methods of stress assessment, such as self-report questionnaires and periodic psychological evaluations, are insufficient for real-time stress monitoring. There is a need for continuous, objective, and non-invasive methods to monitor stress levels, enabling timely interventions to mitigate stress and prevent burnout.

1. **Limitations of Traditional Methods:** Self-report questionnaires are subjective and prone to bias, while periodic evaluations fail to capture the dynamic nature of stress. Research by Stone et al. [36] and Huprich et al. [37] discusses the limitations of these traditional methods.

2. **Advantages of Wearable Technology:** Wearable devices provide continuous monitoring of physiological signals, offering a more accurate and real-time assessment of stress levels. Studies by Samson et al. [38] and Hickey et al. [39] highlight the advantages of using wearables for stress monitoring.

2.3.3 Machine Learning for Enhanced Stress Detection

Using machine learning with data from wearables could help make strong and precise stress detectors. Machine learning can look at big sets of data and find tricky patterns that regular stats might not catch. This could lead to better ways of spotting stress early and helping people manage it well.

1. **Handling Large Datasets:** Wearable devices generate data with higher frequencies, which leads to a large dataset, that can be effectively analyzed using machine learning techniques. Studies by Divya et al. [40] and Sarker et al. [41] discuss the capabilities of machine learning in handling big data.

2. **Capturing Complex Patterns:** Machine learning models, particularly deep learning, can capture complex patterns in physiological signals, enhancing stress classification accuracy. Research by LeCun et al. [42] and Lai et al. [43] showed how powerful deep learning techniques are in tasks like pattern recognition.

2.4 Challenges

2.4.1 Data Quality and Preprocessing

Ensuring the quality and consistency of physiological data collected from wearable devices are the primary challenges while developing a stress classification model. Data preprocessing is crucial for handling noise, artifacts, and missing values.

1. **Noise and Artifacts:** Physiological data collected from wearables can be noisy and contain artifacts due to movement, device placement, and environmental factors. Studies by Charlton et al. [44] and Reiss et al. [45] discuss methods for noise reduction and artifact removal.

2. **Missing Data:** Wearable devices may occasionally fail to record data due to technical issues or user non-compliance. Research by Schafer and Graham [46] and Rubin [47] explores techniques for handling missing data in physiological datasets.

2.4.2 Model Development and Validation

Developing and validating machine learning models for stress classification involves several challenges, including selecting appropriate models, tuning hyperparameters, and ensuring model generalizability.

1. **Model Selection:** Choosing the right machine learning model for stress classification requires evaluating different models' strengths and weaknesses. Studies by Qi et al. [48] and Ding et al. [49] provide guidance on model selection for time-series analysis and classification tasks.

2. **Hyperparameter Tuning:** Improving model performance requires fine-tuning hyperparameters. Some popular hyperparameter tuning techniques include grid search, random search, and Bayesian optimization. Research by Bardenet et al. [50] and Joy et al. [51] discusses various hyperparameter optimization methods.

3. **Generalizability:** Generalizability is important to deploy a model in the real world, and it is a challenge that developed models will produce better results on unseen data. One approach is to test the model with completely new data that was not in the training set. Studies by Chicco D et al. [52] highlight best practices for model validation.

2.4.3 Practical Implementation and Ethical Considerations

Implementing stress monitoring systems in real-world healthcare settings presents practical and ethical challenges, including user acceptance, data privacy, and intervention strategies.

1. **User Acceptance:** For stress monitoring systems to be effective, they must be accepted and regularly used by healthcare professionals. Factors influencing user acceptance include device comfort, ease of use, and perceived benefits. Research by Ma et al. [53] and Durodolu et al. [54] explores the factors affecting technology acceptance.

2. **Data Privacy:** Data collected from the human body as physiological signals raise concerns about data privacy and security. Ensuring that data is anonymized and securely stored is crucial to protecting users' privacy. Studies by Stopczynski et al.

[55] and Filkins et al. [56] discuss best practices for data privacy and security in health monitoring systems.

3. **Intervention Strategies:** Effective intervention strategies are necessary to address detected stress levels. These strategies can include personalized feedback, stress management techniques, and professional support. Research by Giga et al. [57] and Tetrick et al. [58] discusses various intervention approaches for stress management.

CHAPTER 3
RESEARCH METHODOLOGY

3.1 Introduction

This part discusses the research methodology used in developing a stress classification model utilizing physiological data collected from nurses in a hospital environment. The dataset [9], titled "A multimodal sensor dataset for continuous stress detection of nurses in a hospital," provides a comprehensive collection of physiological signals recorded using wearable technology. The methodology encompasses the selection of research subjects, data description, proposed methodology, model workflow, and implementation requirements. We detail the process of developing an LSTM model, including the steps involved in data preprocessing, feature extraction, model training, and evaluation.

3.2 Data Description

The dataset utilized for this study is titled "A multimodal sensor dataset for continuous stress detection of nurses in a hospital," providing a rich collection of physiological signals to study stress in a real-world hospital environment. This dataset is a significant contribution to the research on occupational stress, particularly in the context of nursing, which is a high-stress profession, especially during the COVID-19 pandemic.

3.3 Dataset Overview

The dataset includes physiological data collected from 15 nurses. At the time of the COVID-19 outbreak, they were working in a hospital. According to the authors of the data paper, their goal was to create this dataset to help research understanding and improve the emotional health of nurses by developing early detection methods for work-related stress. The device they have used for the data collection was Empatica E4 smartwatch, which continuously monitors several physiological signals related to stress detection. It has different sensors [6], Photoplethysmography (PPG) to measure Blood Volume Pulse (BVP), electrodermal activity (EDA) sensor to measure change in conductivity in skin, temperature sensor to measure skin temperature, and accelerometer (ACC) sensor to measure physical activity.

The collection is divided into directories and files, with each nurse's information maintained in a distinct folder. Raw data signals are stored in CSV format, in one folder for each participant, and include the following files:

1. **EDA**.csv - Consists of Electrodermal Activity (EDA) signals
2. **HR**.csv - Consists of Heart Rate signals
3. **TEMP**.csv - Consists of Skin Temperature signals
4. **IBI**.csv - Consists of Inter-Beat Interval (IBI) signals
5. **BVP**.csv - Consists of Blood Volume Pulse (BVP) signals
6. **ACC**.csv - Consists of Accelerometer data signals

Each CSV file provides detailed information on the respective physiological signal, including the start time (epoch) and frequency of data collection.

3.4 Physiological Signals

1. **Electrodermal Activity (EDA):** EDA measures the skin's conductivity levels, which vary with the moisture level of the skin, influenced by sweat gland activity. Research has shown that EDA is a good indicator of physiological arousal and stress.

2. **Heart Rate (HR):** HR is not directly collected from the device, it is derived from the other signal, which is the BVP signal. HR is measured in beats per minute (bpm). The HR.csv file contains HR data. Research has shown HR is crucial for understanding the cardiovascular stress response.

3. **Skin Temperature (TEMP):** The skin temperature data, captured in degrees Celsius, provides insights into peripheral blood flow, which can be affected by stress. This data is stored in the file named TEMP.csv.

4. **Inter-Beat Interval (IBI):** The IBI is the time interval between consecutive heartbeats, measured in seconds. This data helps in analyzing heart rate variability (HRV), an important metric for stress analysis. The IBI.csv file contains this data.

5. **Blood Volume Pulse (BVP):** BVP measures blood flow variations in the wrist. From this signal, two other signals are derived, and they are HR and IBI. The BVP.csv file includes this data.

6. **Accelerometer Data (ACC):** The accelerometer data provides information on the physical activity and movement of the nurses, which can contextualize the physiological data. The accelerometer data is measured for the three different axes, x, y, and z, and stored in the ACC.csv file.

3.5 Survey Data

In addition to the physiological data, the dataset includes survey results that capture the context of stressful events. The survey data are stored in an Excel file (SurveyResults.xlsx). The survey captures various aspects of the nurses' stress levels and the contributing factors. The Excel sheets are organized into three groups:

1. **General Information:** Includes columns for anonymized user IDs, event start and end times, duration, and data collection date.

2. **Stress Level:** Mentioned the reported stress level of the nurse during the event.

3. **Stress Contributing Factors:** This covers a wide range of elements, including the stress associated with treating COVID-19 patients, managing patient crises, interacting with patients' families, doctors, and coworkers, administrative duties, an increased workload, stress associated with technology, a shortage of supplies, inadequate documentation, safety concerns, and problems with the work environment.

3.6 Data Collection Procedure

The data were gathered over a week from 15 female nurses, who were working normal shifts at a hospital at the time of COVID-19 outbreak. The age of the participant nurses was between 30 and 55 years. The data collection spanned from April 14, 2020, to December 11, 2020, divided into two phases. To minimize interference with the nurses' activities and adjust the frequency of surveys, the research team carried out a pilot study. Nurses were recruited with the help of the hospital's executive team, nurse managers, and human resources, and they provided consent for their anonymized data to be publicly released.

3.7 Inclusion and Exclusion Criteria

For the data collection, the inclusion and exclusion criteria ensured a representative and healthy sample of the nursing population.

3.7.1 Inclusion Criteria

The inclusion criteria for the study were designed to recruit a specific subset of the nursing staff that would provide relevant data without compromising their health or the study's integrity. These criteria included:

1. **Professional Status:** The study targeted registered nurses actively working in the hospital's Emergency Room (ER) department. This focus on ER nurses was due to the high-stress nature of their work environment, especially during the COVID-19 pandemic.

2. **Gender and Age:** The study included female nurses aged between 30 and 55 years. This age range was chosen to balance the inclusion of both relatively new and more experienced nurses, providing a comprehensive view of stress across different career stages.

3. **Shift Work:** Nurses who worked regular shifts were included, ensuring that the data captured reflected typical work patterns and stressors encountered during their duties. This inclusion criterion was crucial to understanding the day-to-day stress fluctuations experienced by nurses.

3.7.2 Exclusion Criteria

The exclusion criteria were set to ensure the safety of participants and the accuracy of the physiological data collected. These criteria included:

1. **Pregnancy:** Pregnant nurses were excluded due to potential additional stressors and physiological changes associated with pregnancy, which could confound the study's results.

2. **Heavy Smoking:** Nurses who were heavy smokers were excluded because smoking can significantly alter physiological signals, particularly heart rate and skin conductance, which were key metrics in the study.

3. **Mental Disorders:** Nurses with diagnosed mental disorders were excluded to avoid additional variables that could affect stress levels and physiological responses, ensuring a more controlled and interpretable dataset.

4. **Chronic or Cardiovascular Diseases:** Nurses with chronic or cardiovascular diseases were excluded due to the potential impact of these conditions on physiological signals. Chronic illnesses and cardiovascular conditions can alter heart rate, skin temperature, and electrodermal activity, which were critical to the study's measurements.

3.8 Signal Preprocessing

Signal preprocessing is a critical step in preparing the physiological data for model training and analysis. For this study, the raw physiological signals collected using the Empatica E4 wristband, including EDA HR, TEMP, and ACC, were processed to ensure uniformity in frequency, enabling effective integration and analysis.

Initially, all signals were aligned to a frequency of 32 Hz. However, this high frequency resulted in an excessive number of data instances, demanding significant computational resources and memory, which impeded the efficiency of model training. To address this, the signals were down sampled and up sampled to a frequency of 4 Hz, balancing the dataset size and computational feasibility without compromising the integrity of the physiological information.

Specifically, the accelerometer data (ACC) was down sampled from its original higher frequency to 4 Hz, while the heart rate (HR) signal was up sampled to match this frequency. This uniform sampling rate of 4 Hz for all signals was achieved using the SciPy library, leveraging its powerful tools for signal processing. One-dimensional interpolation was employed to up sample the HR signal, ensuring a continuous and smooth representation of the data across time.

This preprocessing strategy effectively harmonized the dataset, making it manageable for machine learning tasks while preserving the essential characteristics of the physiological signals necessary for accurate stress detection.

3.9 Long Short-Term Memory (LSTM) Networks

One of the different types of recurrent neural networks (RNNs) is an LSTM network that is designed in such a way that it can handle sequential data while maintaining long-term dependencies. This capability to remember past inputs for extended periods is crucial for capturing temporal patterns in physiological signals, making LSTM networks particularly useful for time-series data. In tasks like NLP, LSTM is widely used.

3.9.1 LSTM Cell Architecture

An LSTM cell has three gates: the input gate, the forget gate, and the output gate.

1. **Input Gate:** By determining how much fresh data from the data stream should be permitted into the cell's memory, this gate functions as a filter.

2. **Forget Gate:** The forget gate determines which previously recorded information in the cell's memory should be retained and which should be forgotten by concentrating on what is already there.

3. **Output Gate:** The information leaving the cell is managed by this gate. To create the cell's output, it combines the fresh data that goes through the input gate with the pertinent historical data from the forgotten gate.

By working together, these gates address the vanishing gradient problem that plagues traditional RNNs. This allows LSTMs to effectively learn and remember long-term dependencies within data sequences.

3.10 Data Segmentation for LSTM

With a 50% overlap, preprocessed data were segmented with a window size of 10 seconds. Given the 4 Hz sampling rate, each 10-second window comprised 40 data points. Overlapping windows were employed to enhance the temporal resolution and increase the number of training instances without significantly expanding the dataset size.

LSTMs, designed in such a way that they will have the capability to understand temporal dependencies and patterns in sequential data, require windowed segments to effectively learn from the time-series signals. This segmentation ensures that each input sequence retains the temporal context needed for the LSTM to detect stress-related patterns and transitions, thereby improving the model's ability to predict stress accurately.

3.11 Model Development

The architecture consists of three stacked LSTM layers, followed by fully connected dense layers. The detailed model architecture is as follows:

3.11.1 Model Architecture Components

1. **LSTM Layers:** Three LSTM layers, each with 100 units, are used to capture the temporal dependencies in the physiological data. The `return_sequences = True` parameter in the first two LSTM layers ensures that the full sequence is returned, allowing for the stacking of LSTM layers.

2. **Dropout Layers:** After every LSTM layer, dropout layers with a dropout rate of 0.2 are added to avoid overfitting.

3. **Dense Layers:** To add non-linearity, a dense layer with 100 units and a ReLU activation function is introduced after the LSTM and Dense Layers. Another dense layer with a softmax activation function is the output or final layer. The anticipated odds for each class—high stress against no stress, low stress versus no stress—are generated.

3.11.2 Model Compiling

Three LSTM layers make up this model architecture, and to stop overfitting, a dropout layer comes after each LSTM layer. In order to capture intricate correlations in the data, dense layers are built. The metrics, loss function, and optimizer that were employed are broken down as follows:

1. **Optimizer (Adam):** During training, the Adam optimizer dynamically modifies the learning rates according to the gradients' first and second moments. The model is able to converge more quickly and effectively because of this adaptive learning.

2. **Loss Function (Categorical Cross-Entropy):** The loss function that is selected is the categorical cross-entropy. This kind of multiclass classification task is appropriate for it. It calculates the discrepancy between each class's actual labels and projected probability. In order to make sure that the model's predictions closely match the actual class labels, the goal is to minimize this difference.

3. **Metrics (Accuracy):** In training and testing, accuracy serves as the evaluation metric. Out of all cases, it determines the percentage of correctly classified instances. Stated differently, it indicates the degree to which the model's predictions agree with the actual class designations. Higher accuracy shows that the model performed better at classifying the data into the appropriate groups.

The model's performance depends on the optimizer and loss function selection because they affect the training dynamics and convergence rate.

3.12 Performance Metrics

To evaluate the performance of the LSTM model in classifying stress levels, several key metrics are used: accuracy, precision, recall, F1-score, and the confusion matrix. Each metric provides unique insights into the model’s performance and helps in understanding its effectiveness in different aspects.

3.12.1 Accuracy

Accuracy is the simplest way to measure a model's performance. It shows the percentage of correct predictions out of all predictions made. In other words, it tells us how often the model gets things right. For example, if a model has 90% accuracy, it means that 90 out of 100 predictions are correct. This gives a good overall idea of how well the model is doing at making predictions.

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

Where:

- **TP** = True Positives
- **TN** = True Negatives
- **FP** = False Positives

·　**FN** = False Negatives

3.12.2 Precision

Precision, also known as Positive Predictive Value, measures the proportion of positive predictions that are correct. It tells us how many of the predicted positive cases actually belong to the positive class. This metric is especially important in cases where the cost of false positives is high, meaning that it helps to avoid errors where something is wrongly predicted as positive. High precision ensures that most positive predictions made by the model are accurate and trustworthy. And the formula to calculate precision is as follows:

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

3.12.3 Recall

Recall, also known as Sensitivity, measures how many of the actual instances of stress the model correctly identifies. It indicates the model's capability to capture all cases of stress accurately. This metric is particularly vital in stress monitoring scenarios, where it's crucial not to overlook any instances of stress. Ensuring high recall helps in promptly identifying and addressing stress levels among individuals, thereby supporting their well-being and maintaining a healthy work environment. And the formula is as follows:

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

3.12.4 F1-Score

The F1-score integrates precision and recall into a single metric, achieving a balance between them. This single score provides a comprehensive measure of the model's performance, taking into account both the accuracy with which it finds positive situations and its ability to avoid false positives. It's especially useful when dealing with unequal class distributions, in which one class is significantly more prevalent than others. This statistic ensures a fair evaluation of the model's efficacy across all classes, independent of distribution.

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

3.12.5 Confusion Matrix

The confusion matrix is a tabular representation of the actual versus predicted classifications, allowing for a detailed analysis of the model's performance. It includes the counts of true positives, true negatives, false positives, and false negatives.

- **TP** (True Positives): Correctly predicted positive instances.
- **TN** (True Negatives): Correctly predicted negative instances.
- **FP** (False Positives): Incorrectly predicted positive instances.
- **FN** (False Negatives): Incorrectly predicted negative instances.

By utilizing these performance metrics, the study provides a comprehensive evaluation of the LSTM model's ability to classify stress levels. Accuracy offers an overall performance measure, precision and recall provide insights into the handling of positive instances, the F1-score balances these aspects, and the confusion matrix offers a detailed view of prediction outcomes.

3.13 Implementation Requirements

Implementing the proposed LSTM model for stress classification requires specific hardware and software resources. The following details outline the implementation requirements:

3.13.1 Hardware Requirements

1. **Computational Resources:** High-performance computing resources, such as GPUs (Graphics Processing Units), are recommended for training the LSTM model due to their ability to handle the computational demands of deep learning. GPUs accelerate the matrix operations involved in training, significantly reducing the time required.

2. **Wearable Devices:** Empatica E4 smartwatches for data collection from participants, ensuring continuous monitoring of physiological signals.

3.13.2 Software Requirements

1. **Programming Language:** Python, due to its extensive libraries and frameworks for machine learning and data analysis.

2. **Deep Learning Frameworks:** The LSTM model was constructed and trained using TensorFlow and Keras. High-level APIs offered by these frameworks make it easier to create complicated neural networks.

3. **Data Processing Libraries:** NumPy, Pandas, and Scikit-learn for data preprocessing and feature extraction. These libraries offer efficient tools for handling large datasets and performing data transformations.

4. **Development Environment:** The development environment for this study utilizes Jupyter Notebook for interactive coding and visualization.

CHAPTER 4
EXPERIMENTAL RESULTS AND DISCUSSION

4.1 Introduction

The experimental results and discussion section presents the outcomes of the developed LSTM models for stress classification using physiological signals from the Empatica E4 smartwatch. This section includes a detailed analysis of the models' performance on both binary and multiclass classification tasks. By comparing these results, we gain insights into the effectiveness of the models and the challenges associated with each classification approach.

4.2 Experimental Results

4.2.1 Binary Classification

For the binary classification task, the stress levels were simplified into two categories: stress (merging low stress and high stress) and no stress. The classification report for the binary model is shown in Table 1.

Table 1: Binary Classification Report

Class	Precision	Recall	F1-Score	Support
0 (No Stress)	0.90	0.82	0.86	2576
2 (Stress)	0.91	0.96	0.93	4942
Accuracy	0.91			7518

The binary classification model also achieved an overall accuracy of 0.91. The performance metrics indicate a high precision and recall for both classes, with particularly strong performance in detecting stress.

4.2.2 Multiclass Classification

In the multiclass classification task, the model distinguishes between three stress levels: no stress, low stress, and high stress. The classification report for this model is summarized in Table 2.

Table 2: Multiclass Classification Report

Class	Precision	Recall	F1-Score	Support
0 (No Stress)	0.90	0.82	0.86	2136
1 (Low Stress)	0.74	0.71	0.73	440
2 (High Stress)	0.94	0.94	0.94	4942
Accuracy	0.91			7518

The results indicate high accuracy (0.91) across the multiclass classification task. The precision, recall, and F1-score for each class show that the model performs exceptionally well in identifying "no stress" and "high stress" states, with slightly lower performance in detecting "low stress."

4.3 Discussion

4.3.1 Multiclass vs. Binary Classification

The experimental results demonstrate that the LSTM model performs well in both binary and multiclass classification tasks, achieving an accuracy of 0.91 in both cases. However, there are notable differences in performance metrics for specific classes.

4.3.1.1 Multiclass Classification

In the multiclass classification task, the model shows excellent performance in detecting "no stress" (precision: 0.88, recall: 0.90) and "high stress" (precision: 0.94, recall: 0.94). This indicates that the physiological signals captured by the Empatica E4 smartwatch are effective in distinguishing between states of no stress and high stress.

However, the performance drops for the "low stress" category (precision: 0.74, recall: 0.71). This could be due to several factors:

1. **Ambiguity in Low Stress Signals:** The physiological signals for low stress might overlap significantly with those of no stress and high stress, making it challenging for the model to distinguish between these states accurately.
2. **Class Imbalance:** The support for the low stress category (440 instances) is significantly lower compared to the no stress (2136 instances) and high stress (4942 instances) categories. This imbalance can lead to biased learning where the model is less sensitive to the less frequent class.

4.3.1.2 Binary Classification

In the binary classification task, the model maintains high precision and recall for both classes. The model's performance in detecting stress (precision: 0.91, recall: 0.96) is particularly noteworthy, indicating a strong ability to identify stressful states. The no stress class also shows robust performance (precision: 0.90, recall: 0.82).

The binary classification results suggest that merging low stress with no stress simplifies the classification task, reducing the overlap and ambiguity present in the multiclass scenario. This simplification leads to more consistent performance across both classes.

4.3.2 Interpretation of Metrics

1. **Precision:** High precision in both models means that when the model predicts a class (like stress or no stress), it's usually right. This is important in healthcare because false positives, like wrongly predicting high stress, can lead to unnecessary actions.

2. **Recall:** High recall, especially for detecting high stress in both binary and multiclass models, means the model is good at finding real cases of stress. This is crucial for timely and appropriate actions in hospitals.

3. **F1-Score:** The F1-score combines both precision and recall into one number. High F1-scores across all classes show that the model is reliable for classifying stress accurately.

4.3.3 Confusion Matrix Analysis

To further analyze the model's performance, confusion matrices for both binary and multiclass classification are presented in Figure 1 and 2.

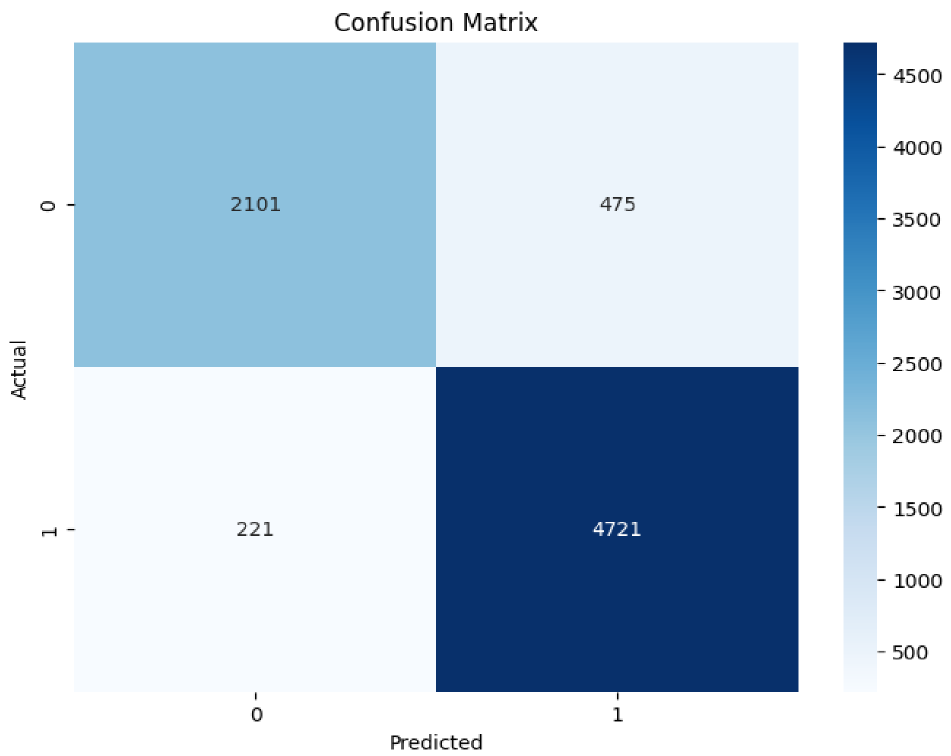

Figure 1: Confusion Matrix for Binary Classification

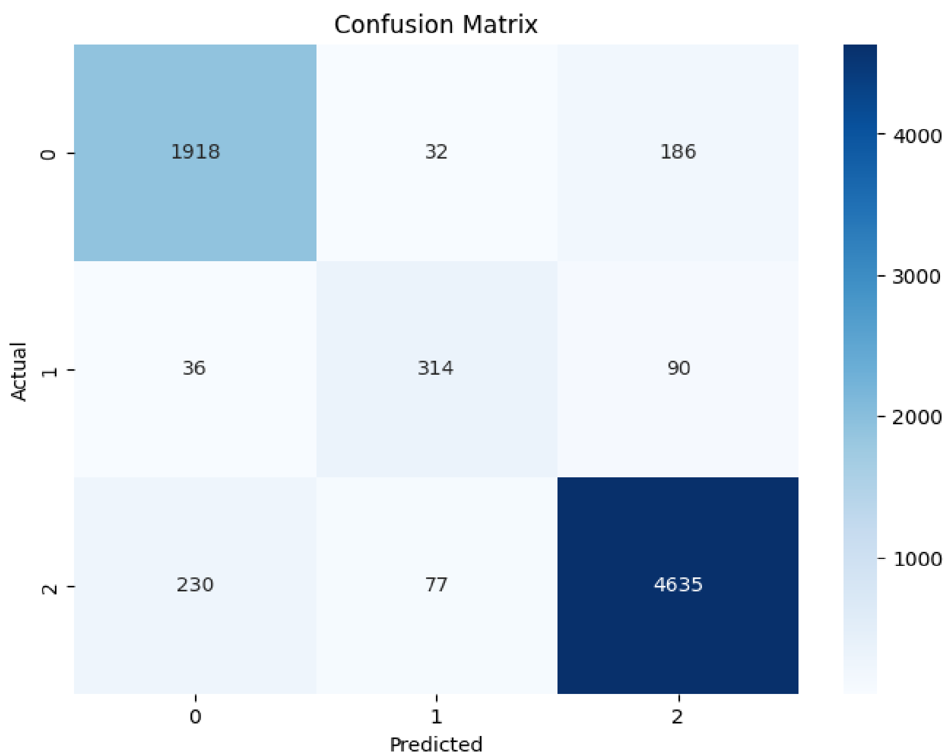

Figure 2: Confusion Matrix for Multiclass Classification

The confusion matrices show a detailed view of the model’s performance. In the multiclass scenario, misclassifications between no stress and low stress are relatively common, as seen from the off-diagonal elements. In comparison, the binary classification matrix has fewer misclassifications, demonstrating the model's ability to differentiate between stress and no stress more efficiently when the categories are combined.

The experimental results demonstrate that the LSTM model is highly effective in classifying stress levels using physiological signals. While both binary and multiclass models achieve high accuracy, the binary classification model exhibits slightly better performance in distinguishing between stress and no stress. The complete study of performance indicators, including accuracy, precision, recall, F1-score, and confusion matrices, provides a thorough understanding of the model's strengths and weaknesses.

Future work could explore techniques to improve the detection of low stress in the multiclass model, such as data augmentation to address class imbalance or advanced signal processing methods to enhance feature extraction. Additionally, further validation in real-world hospital settings would help assess the model's practical utility and robustness.

CHAPTER 5
IMPACT ON SOCIETY, AND SUSTAINABILITY

5.1 Impact on Society

The implementation of an LSTM model for stress classification among nurses in a hospital environment has significant societal implications. By accurately identifying stress levels, hospital administrators and healthcare providers can take proactive steps to manage workplace stress, which has profound effects on the well-being of healthcare professionals and the quality of patient care.

5.1.1 Improving Nurse Well-being

Nurses are the backbone of the healthcare system, and their well-being directly impacts their ability to provide high-quality care. Chronic stress can lead to burnout, reduced job satisfaction, and increased turnover rates, all of which negatively affect the healthcare system. The ability to monitor and address stress in real-time enables hospitals to provide timely interventions, such as breaks, counseling, or workload adjustments. This support not only improves nurses' mental and physical health but also enhances their job satisfaction and retention.

5.1.2 Enhancing Patient Care

Stress among nurses can compromise patient care, as it affects their ability to concentrate, make decisions, and interact compassionately with patients. By implementing stress management strategies based on real-time data, hospitals can ensure that nurses are in optimal condition to perform their duties. This leads to better patient outcomes, reduced medical errors, and a higher overall quality of care.

5.1.3 Promoting a Supportive Work Environment

The adoption of stress monitoring technology fosters a culture of support and empathy within healthcare institutions. It signals to employees that their well-being is a priority, which can enhance morale and foster a sense of community. This supportive environment is crucial for attracting and retaining skilled healthcare professionals, especially in an industry facing workforce shortages.

5.2 Sustainability Plan

A sustainability plan ensures the long-term viability and positive impact of stress monitoring systems in healthcare environments. This involves strategic planning, resource allocation, and continuous evaluation.

5.2.1 Long-Term Monitoring and Support

Implementing stress monitoring systems should not be a one-time initiative but part of an ongoing commitment to employee well-being. Hospitals should establish long-term monitoring programs with regular assessments to ensure they continue to meet the needs of nurses. Continuous support mechanisms, such as access to mental health resources and stress management training, should be integrated into the program.

CHAPTER 6

CONCLUSION AND FUTURE WORK

6.1 Summary of the Study

The primary focus of this study was to develop and evaluate an LSTM model for stress classification among nurses in a hospital environment using physiological data collected via the Empatica E4 smartwatch. The overarching goal was to provide a tool that could assist hospital management in identifying and managing workplace stress among nursing staff. Stress was classified into three levels based on survey results (no stress, low stress, and high stress), and two distinct models were developed using extensive data processing due to the time series nature of the data: a binary classification model and a multiclass classification model.

The dataset utilized was "A multimodal sensor dataset for continuous stress detection of nurses in a hospital," which included various physiological signals such as electrodermal activity (EDA), heart rate (HR), accelerometer (ACC) data, and temperature (TEMP). The preprocessing steps involved down sampling and up sampling these signals to a consistent frequency of 4 Hz, followed by segmenting the data into 10-second windows with a 50% overlap to accommodate the LSTM architecture.

The LSTM model was built with three layers of LSTM units stacked on top of each other. Each of these LSTM layers was followed by a dropout layer to help prevent overfitting, which means the model won't just memorize the training data. Dense layers were added to capture complex patterns in the data. The model was compiled using the Adam optimizer and a categorical cross-entropy loss function. We evaluated the model's performance using metrics like accuracy, precision, recall, F1-score, and confusion matrices to ensure it worked well.

6.2 Conclusion

The experimental results demonstrated that the LSTM models were effective in classifying stress levels among nurses, with both binary and multiclass models achieving an accuracy of 0.91. The detailed performance metrics for each model provided insights into their strengths and limitations:

- **Multiclass Classification:** The multiclass model performed exceptionally well in identifying "no stress" and "high stress" states, with precision and recall values exceeding 0.88. However, the model showed slightly lower performance in detecting "low stress," likely due to the overlapping nature of physiological signals and class imbalance.

- **Binary Classification:** The binary model, which merged low stress with no stress, exhibited high precision and recall for both stress and no stress categories. This simplification improved the model's ability to distinguish between stress and no stress, highlighting the potential benefits of reducing the complexity of classification tasks.

The successful use of these models shows that it's possible to use wearable technology and advanced machine learning to monitor and manage workplace stress in real-time. This means that we can keep track of stress levels as they happen, which is very helpful. By understanding stress levels better, hospitals can take action to help nurses manage their stress. This support can improve nurses' well-being and make them feel better at work. As a result, patient care can improve, and the work environment can become healthier and more supportive. Overall, using these tools can help create a better, less stressful place for nurses to work and care for patients.

6.3 Implications for Further Study

This study has shown that LSTM models can be useful for classifying stress. However, there are many ways to improve and expand this approach in future research to make it more effective and widely applicable.

6.3.1 Advanced Signal Processing Techniques

Future research could explore better ways to process signals from physiological data to improve how we extract important features. For example, using wavelet transforms, principal component analysis (PCA), or deep learning techniques could help find more detailed patterns in the data. This could make the models more accurate, especially in identifying low levels of stress that might be missed otherwise. By enhancing these techniques, we could improve the performance of stress detection models and make them more reliable and sensitive to subtle changes in stress levels. This would help in creating better tools for monitoring and managing stress in various settings.

6.3.2 Addressing Class Imbalance

Class imbalance was identified as a challenge in our multiclass classification task. This means that some stress levels were much less represented in the data compared to others, making it harder for the model to learn and classify these less common classes accurately. To address this issue, techniques like data augmentation and synthetic data generation, such as using T-SMOTE (Temporal-oriented Synthetic Minority Over-sampling Technique) [59], can be used to create more balanced datasets. By ensuring a more balanced representation of stress levels in the data, we can help the model learn to classify all stress categories more accurately and fairly, improving its overall performance and reliability. This will lead to a better understanding and detection of different levels of stress in real-world applications.

6.3.3 Real-world Validation and Integration

While the study's results look promising, it's important to test the models in real-life hospital settings to see how useful they really are. To do this, we could run pilot studies where the stress monitoring system is used in hospitals as part of daily routines. Feedback from nurses and hospital management would be valuable to improve and fine-tune the system. Additionally, it's important to see how the stress monitoring system can work with existing hospital information systems and workflows, making it easier to adopt. By doing these things, we can ensure that the system is practical, effective, and easy to use in real hospital environments, helping to monitor and manage stress more efficiently.

6.3.4 Longitudinal Studies and Chronic Stress Monitoring

This study focused on real-time stress monitoring, but future research could extend this to longitudinal studies that track stress levels over longer periods. This would provide insights into chronic stress patterns and their impact on nurses' health and job performance. Understanding long-term stress dynamics could inform more comprehensive stress management strategies and interventions.

6.3.5 Expanding to Other Healthcare Professionals

While this study focused on nurses, other healthcare professionals, such as doctors, technicians, and administrative staff, also experience significant workplace stress. Expanding the research to include these groups could provide a more holistic understanding of stress in healthcare settings and lead to the development of comprehensive stress management programs for the entire workforce.

6.3.6 Incorporating Psychological and Environmental Factors

Physiological signals provide valuable insights into stress, but incorporating psychological assessments and environmental factors could enhance the accuracy and context of stress classification. Future studies could integrate self-reported stress measures, work environment data, and social interactions to develop a more comprehensive stress monitoring system.

REFERENCE

[1] S. J. Lupien, F. Maheu, M. Tu, A. Fiocco, and T. E. Schramek, “The effects of stress and stress hormones on human cognition: Implications for the field of brain and cognition,” Brain and Cognition, vol. 65, no. 3, pp. 209–237, Dec. 2007, doi: https://doi.org/10.1016/j.bandc.2007.02.007.

[2] J. Axelrod and T. Reisine, “Stress hormones: their interaction and regulation,” Science, vol. 224, no. 4648, pp. 452–459, May 1984, doi: https://doi.org/10.1126/science.6143403.

[3] S. A. Adamo, “The Effects of Stress Hormones on Immune Function May be Vital for the Adaptive Reconfiguration of the Immune System During Fight-or-Flight Behavior,” Integrative and Comparative Biology, vol. 54, no. 3, pp. 419–426, Mar. 2014, doi: https://doi.org/10.1093/icb/icu005.

[4] A.-R. Babapour, N. Gahassab-Mozaffari, and A. Fathnezhad-Kazemi, “Nurses’ Job Stress and Its Impact on Quality of Life and Caring behaviors: a cross-sectional Study,” BMC Nursing, vol. 21, no. 1, pp. 1–10, Mar. 2022, Available: https://bmcnurs.biomedcentral.com/articles/10.1186/s12912-022-00852-y

[5] C. Demetriou, B. U. Ozer, and C. A. Essau, “Self-Report Questionnaires,” The Encyclopedia of Clinical Psychology, pp. 1–6, Jan. 2015, doi: https://doi.org/10.1002/9781118625392.wbecp507.

[6] A. A. T. Schuurmans et al., “Validity of the Empatica E4 Wristband to Measure Heart Rate Variability (HRV) Parameters: a Comparison to Electrocardiography (ECG),” Journal of Medical Systems, vol. 44, no. 11, Sep. 2020, doi: https://doi.org/10.1007/s10916-020-01648-w.

[7] Y. S. Can, B. Arnrich, and C. Ersoy, “Stress detection in daily life scenarios using smart phones and wearable sensors: A survey,” Journal of Biomedical Informatics, vol. 92, p. 103139, Apr. 2019, doi: https://doi.org/10.1016/j.jbi.2019.103139.

[8] S. Hochreiter and J. Schmidhuber, “Long Short-Term Memory,” Neural Computation, vol. 9, no. 8, pp. 1735–1780, Nov. 1997, doi: https://doi.org/10.1162/neco.1997.9.8.1735.

[9] S. Hosseini et al., “A multimodal sensor dataset for continuous stress detection of nurses in a hospital,” Scientific Data, vol. 9, no. 1, Jun. 2022, doi: https://doi.org/10.1038/s41597-022-01361-y.

[10] P. Bobade and M. Vani, “Stress Detection with Machine Learning and Deep Learning using Multimodal Physiological Data,” 2020 Second International Conference on Inventive Research in Computing Applications (ICIRCA), Jul. 2020, doi: https://doi.org/10.1109/icirca48905.2020.9183244.
[11] P. Schmidt, A. Reiss, R. Duerichen, C. Marberger, and K. Van Laerhoven, “Introducing WESAD, a Multimodal Dataset for Wearable Stress and Affect Detection,” Proceedings of the 20th ACM International Conference on Multimodal Interaction, Oct. 2018, doi: https://doi.org/10.1145/3242969.3242985.
[12] Z. L, “PRIME PubMed | Stress Detection Through Wrist-Based Electrodermal Activity Monitoring and Machine Learning,” www.unboundmedicine.com, May 2023. https://www.unboundmedicine.com/medline/citation/37022004/Stress_Detection_Through_Wrist_Based_Electrodermal_Activity_Monitoring_and_Machine_Learning_ (accessed Jul. 03, 2024).
[13] Z. Z, “Stress Detection Using Machine Learning and Deep L | Download Free PDF | Support Vector Machine | Machine Learning,” Scribd, 2021. https://www.scribd.com/document/655472860/Stress-Detection-Using-Machine-Learning-and-Deep-L (accessed Jul. 03, 2024).
[14] R. K. Nath, H. Thapliyal, and A. Caban-Holt, “Machine Learning Based Stress Monitoring in Older Adults Using Wearable Sensors and Cortisol as Stress Biomarker,” Journal of Signal Processing Systems, Jan. 2021, doi: https://doi.org/10.1007/s11265-020-01611-5.
[15] R. Tanwar, O. Phukan, G. Singh, and S. Tiwari, “CNN-LSTM Based Stress Recognition Using Wearables,” 2022. Accessed: May 14, 2023. [Online]. Available: https://ceur-ws.org/Vol-3335/DLQ_short2.pdf
[16] B. Roy et al., “Hybrid Deep Learning Approach for Stress Detection Using Decomposed EEG Signals,” Diagnostics, vol. 13, no. 11, pp. 1936–1936, Jun. 2023, doi: https://doi.org/10.3390/diagnostics13111936.
[17] J. S. Banerjee, M. Mahmud, and D. Brown, “Heart Rate Variability-Based Mental Stress Detection: An Explainable Machine Learning Approach,” SN Computer Science, vol. 4, no. 2, Jan. 2023, doi: https://doi.org/10.1007/s42979-022-01605-z.
[18] K. M. Dalmeida and G. L. Masala, “HRV Features as Viable Physiological Markers for Stress Detection Using Wearable Devices,” Sensors, vol. 21, no. 8, p. 2873, Apr. 2021, doi: https://doi.org/10.3390/s21082873.

[19] F. Vasile, A. Vizziello, Natascia Brondino, and Pietro Savazzi, “Stress State Classification Based on Deep Neural Network and Electrodermal Activity Modeling,” Sensors, vol. 23, no. 5, pp. 2504–2504, Feb. 2023, doi: https://doi.org/10.3390/s23052504.
[20] P. Karthikeyan, M. Murugappan, and S. Yaacob, “Descriptive Analysis of Skin Temperature Variability of Sympathetic Nervous System Activity in Stress,” Journal of Physical Therapy Science, vol. 24, no. 12, pp. 1341–1344, 2012, doi: https://doi.org/10.1589/jpts.24.1341.
[21] C. H. Vinkers et al., “The effect of stress on core and peripheral body temperature in humans,” Stress, vol. 16, no. 5, pp. 520–530, Jul. 2013, doi: https://doi.org/10.3109/10253890.2013.807243.
[22] S. Campanella, A. Altaleb, A. Belli, P. Pierleoni, and L. Palma, “A Method for Stress Detection Using Empatica E4 Bracelet and Machine-Learning Techniques,” Sensors, vol. 23, no. 7, p. 3565, Mar. 2023, doi: https://doi.org/10.3390/s23073565.
[23] G. Y. Lui, D. Loughnane, C. Polley, T. Jayarathna, and P. P. Breen, “The Apple Watch for Monitoring Mental Health–Related Physiological Symptoms: Literature Review,” JMIR Mental Health, vol. 9, no. 9, p. e37354, Sep. 2022, doi: https://doi.org/10.2196/37354.
[24] S. Sharma, G. Singh, and M. Sharma, “A comprehensive review and analysis of supervised-learning and soft computing techniques for stress diagnosis in humans,” Computers in Biology and Medicine, vol. 134, p. 104450, Jul. 2021, doi: https://doi.org/10.1016/j.compbiomed.2021.104450.
[25] N. Phutela, D. Relan, G. Gabrani, P. Kumaraguru, and M. Samuel, “Stress Classification Using Brain Signals Based on LSTM Network,” Computational Intelligence and Neuroscience, vol. 2022, pp. 1–13, Apr. 2022, doi: https://doi.org/10.1155/2022/7607592.
[26] Ankita, S. Rani, Ali Kashif Bashir, Adi Alhudhaif, Deepika Koundal, and Adi Alhudhaif, “An efficient CNN-LSTM model for sentiment detection in #BlackLivesMatter,” vol. 193, pp. 116256–116256, May 2022, doi: https://doi.org/10.1016/j.eswa.2021.116256.
[27] M. Kang, S. Shin, J. Jung, and Y. T. Kim, “Classification of Mental Stress Using CNN-LSTM Algorithms with Electrocardiogram Signals,” Journal of Healthcare Engineering, vol. 2021, p. e9951905, Jun. 2021, doi: https://doi.org/10.1155/2021/9951905.

[28] Y. Acikmese and S. E. Alptekin, “Prediction of stress levels with LSTM and passive mobile sensors,” Procedia Computer Science, vol. 159, pp. 658–667, 2019, doi: https://doi.org/10.1016/j.procs.2019.09.221.

[29] S. Ito, S. Fujita, K. Seto, T. Kitazawa, K. Matsumoto, and T. Hasegawa, “Occupational stress among healthcare workers in Japan,” Work, vol. 49, no. 2, pp. 225–234, 2014, doi: https://doi.org/10.3233/wor-131656.

[30] A. Mohanty, A. Kabi, and A. P. Mohanty, “Health problems in healthcare workers: A review,” Journal of Family Medicine and Primary Care, vol. 8, no. 8, pp. 2568–2572, Aug. 2019, Available: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6753812/

[31] Williams, Eric S., “Handbook of Human Factors and Ergonomics in Health Care and Patient Safety,” Routledge & CRC Press, 2012. https://www.routledge.com/Handbook-of-Human-Factors-and-Ergonomics-in-Health-Care-and-Patient-Safety/Carayon/p/book/9781138074590

[32] N. Khamisa, B. Oldenburg, K. Peltzer, and D. Ilic, “Work Related Stress, Burnout, Job Satisfaction and General Health of Nurses,” International Journal of Environmental Research and Public Health, vol. 12, no. 1, pp. 652–666, Jan. 2015, doi: https://doi.org/10.3390/ijerph120100652.

[33] M. Selamu, G. Thornicroft, A. Fekadu, and C. Hanlon, “Conceptualisation of job-related wellbeing, stress and burnout among healthcare workers in rural Ethiopia: a qualitative study,” BMC Health Services Research, vol. 17, no. 1, Jun. 2017, doi: https://doi.org/10.1186/s12913-017-2370-5.

[34] K. Joshi, B. Modi, S. Singhal, and S. Gupta, Occupational Stress among Health Care Workers. IntechOpen, 2022. Available: https://www.intechopen.com/chapters/83965

[35] H. Wang et al., “Healthcare workers’ stress when caring for COVID-19 patients: An altruistic perspective,” Nursing Ethics, vol. 27, no. 7, p. 096973302093414, Jul. 2020, doi: https://doi.org/10.1177/0969733020934146.

[36] A. A. Stone, M. A. Greenberg, E. Kennedy-Moore, and M. G. Newman, “Self-report, situation-specific coping questionnaires: What are they measuring?,” Journal of Personality and Social Psychology, vol. 61, no. 4, pp. 648–658, 1991, doi: https://doi.org/10.1037/0022-3514.61.4.648.

[37] S. K. Huprich, R. F. Bornstein, and T. A. Schmitt, “Self-Report Methodology is Insufficient for Improving the Assessment and Classification of Axis II Personality

Disorders," Journal of Personality Disorders, vol. 25, no. 5, pp. 557–570, Oct. 2011, doi: https://doi.org/10.1521/pedi.2011.25.5.557.

[38] C. Samson and A. Koh, "Stress Monitoring and Recent Advancements in Wearable Biosensors," Frontiers in Bioengineering and Biotechnology, vol. 8, Sep. 2020, doi: https://doi.org/10.3389/fbioe.2020.01037.

[39] B. A. Hickey et al., "Smart Devices and Wearable Technologies to Detect and Monitor Mental Health Conditions and Stress: A Systematic Review," Sensors, vol. 21, no. 10, p. 3461, May 2021, doi: https://doi.org/10.3390/s21103461.

[40] K Sree Divya, Peyakunta Bhargavi, S Jyothi, "Machine learning algorithms in big data analytics," scholar.google.co.in, 2018. https://scholar.google.co.in/citations?view_op=view_citation&hl=en&user=Nn_JOeYAAAAJ&citation_for_view=Nn_JOeYAAAAJ:ns9cj8rnVeAC (accessed Jun. 07, 2024).

[41] I. H. Sarker, "Machine Learning: Algorithms, Real-World Applications and Research Directions," SN Computer Science, vol. 2, no. 3, pp. 1–21, Mar. 2021, doi: https://doi.org/10.1007/s42979-021-00592-x.

[42] Y. LeCun, "The Power and Limits of Deep Learning," Research-Technology Management, vol. 61, no. 6, pp. 22–27, Nov. 2018, doi: https://doi.org/10.1080/08956308.2018.1516928.

[43] G. Lai, W.-C. Chang, Y. Yang, and H. Liu, "Modeling Long- and Short-Term Temporal Patterns with Deep Neural Networks," arXiv (Cornell University), Mar. 2017, doi: https://doi.org/10.48550/arxiv.1703.07015.

[44] P. H. Charlton et al., "Extraction of respiratory signals from the electrocardiogram and photoplethysmogram: technical and physiological determinants," Physiological Measurement, vol. 38, no. 5, pp. 669–690, Apr. 2017, doi: https://doi.org/10.1088/1361-6579/aa670e.

[45] Lina, R. Ramachandran, and B. J. May, "Effects of Signal Level and Background Noise on Spectral Representations in the Auditory Nerve of the Domestic Cat," Journal of the Association for Research in Otolaryngology, vol. 12, no. 1, pp. 71–88, Sep. 2010, doi: https://doi.org/10.1007/s10162-010-0232-5.

[46] J. Brick and G. Kalton, "Handling missing data in survey research," Statistical Methods in Medical Research, vol. 5, no. 3, pp. 215–238, Sep. 1996, doi: https://doi.org/10.1177/096228029600500302.

[47] L. H. Rubin, K. Witkiewitz, J. St. Andre, and S. Reilly, "Methods for Handling Missing Data in the Behavioral Neurosciences: Don't Throw the Baby Rat out with the Bath Water," Journal of Undergraduate Neuroscience Education, vol. 5, no. 2, pp. A71–A77, Jun. 2007, Available: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3592650/
[48] M. Qi and G. P. Zhang, "An investigation of model selection criteria for neural network time series forecasting," European Journal of Operational Research, vol. 132, no. 3, pp. 666–680, Aug. 2001, doi: https://doi.org/10.1016/s0377-2217(00)00171-5.
[49] J. Ding, V. Tarokh, and Y. Yang, "Model Selection Techniques: An Overview," IEEE Signal Processing Magazine, vol. 35, no. 6, pp. 16–34, Nov. 2018, doi: https://doi.org/10.1109/msp.2018.2867638.
[50] Rémi Bardenet, M. Brendel, Bal, and Mich le Sebag, "Collaborative hyperparameter tuning," International Conference on Machine Learning, vol. 28, pp. 199–207, Jun. 2013.
[51] T. T. Joy, S. Rana, S. Gupta, and S. Venkatesh, "Hyperparameter tuning for big data using Bayesian optimisation," IEEE Xplore, Dec. 01, 2016. https://ieeexplore.ieee.org/abstract/document/7900023
[52] D. Chicco and G. Jurman, "The ABC recommendations for validation of supervised machine learning results in biomedical sciences," Frontiers in Big Data, vol. 5, Sep. 2022, doi: https://doi.org/10.3389/fdata.2022.979465.
[53] Q. Ma and L. Liu, "The Technology Acceptance Model: A Meta-Analysis of Empirical Findings," www.igi-global.com, 2005. https://www.igi-global.com/chapter/technology-acceptance-model/4475 (accessed Jun. 07, 2024).
[54] Durodolu and Oluwole Olumide, "Technology Acceptance Model as a predictor of using information system' to acquire information literacy skills," Nov. 2016.
[55] A. Stopczynski, R. Pietri, A. Pentland, D. Lazer, and S. Lehmann, "Privacy in Sensor-Driven Human Data Collection: A Guide for Practitioners," arXiv (Cornell University), Mar. 2014.
[56] B. L. Filkins et al., "Privacy and security in the era of digital health: what should translational researchers know and do about it?," American journal of translational research, vol. 8, no. 3, pp. 1560–80, 2016, Available: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4859641/
[57] S. I. Giga, C. L. Cooper, and B. Faragher, "The development of a framework for a comprehensive approach to stress management interventions at work.," International

Journal of Stress Management, vol. 10, no. 4, pp. 280–296, 2003, doi: https://doi.org/10.1037/1072-5245.10.4.280.

[58] L. E. Tetrick and C. J. Winslow, “Workplace Stress Management Interventions and Health Promotion,” Annual Review of Organizational Psychology and Organizational Behavior, vol. 2, no. 1, pp. 583–603, Apr. 2015, doi: https://doi.org/10.1146/annurev-orgpsych-032414-111341.

[59] P. Zhao et al., “T-SMOTE: Temporal-oriented Synthetic Minority Oversampling Technique for Imbalanced Time Series Classification.” Accessed: Jun. 07, 2024. [Online]. Available: https://www.ijcai.org/proceedings/2022/0334.pdf

PLAGIARISM REPORT

Watch

ORIGINALITY REPORT

2%	2%	2%	1%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	fastercapital.com Internet Source	2%
2	Sayandeep Ghosh, SeongKi Kim, Muhammad Fazal Ijaz, Pawan Kumar Singh, Mufti Mahmud. "Classification of Mental Stress from Wearable Physiological Sensors Using Image-Encoding-Based Deep Neural Network", Biosensors, 2022 Publication	1%

Exclude quotes	On	Exclude matches	< 1%
Exclude bibliography	Off		