2021

Novel Machine Learning and Wearable Sensor Based Solutions for Smart Healthcare Monitoring

Rajdeep Kumar Nath
University of Kentucky, rajdeepkumar.nath@gmail.com
Digital Object Identifier: https://doi.org/10.13023/etd.2021.254

STUDENT AGREEMENT:

I represent that my thesis or dissertation and abstract are my original work. Proper attribution has been given to all outside sources. I understand that I am solely responsible for obtaining any needed copyright permissions. I have obtained needed written permission statement(s) from the owner(s) of each third-party copyrighted matter to be included in my work, allowing electronic distribution (if such use is not permitted by the fair use doctrine) which will be submitted to UKnowledge as Additional File.

I hereby grant to The University of Kentucky and its agents the irrevocable, non-exclusive, and royalty-free license to archive and make accessible my work in whole or in part in all forms of media, now or hereafter known. I agree that the document mentioned above may be made available immediately for worldwide access unless an embargo applies.

I retain all other ownership rights to the copyright of my work. I also retain the right to use in future works (such as articles or books) all or part of my work. I understand that I am free to register the copyright to my work.

REVIEW, APPROVAL AND ACCEPTANCE

The document mentioned above has been reviewed and accepted by the student's advisor, on behalf of the advisory committee, and by the Director of Graduate Studies (DGS), on behalf of the program; we verify that this is the final, approved version of the student's thesis including all changes required by the advisory committee. The undersigned agree to abide by the statements above.

Rajdeep Kumar Nath, Student

Dr. Himanshu Thapliyal, Major Professor

Dr. Daniel Lau, Director of Graduate Studies

NOVEL MACHINE LEARNING AND WEARABLE SENSOR BASED SOLUTIONS FOR SMART HEALTHCARE MONITORING

DISSERTATION

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
in the College of Engineering
at the University of Kentucky

By
Rajdeep Kumar Nath
Lexington, Kentucky
Director: Dr. Himanshu Thapliyal, Associate Professor of
Electrical and Computer Engineering
Lexington, Kentucky
2021

ABSTRACT OF DISSERTATION

NOVEL MACHINE LEARNING AND WEARABLE SENSOR BASED SOLUTIONS FOR SMART HEALTHCARE MONITORING

The advent of IoT has enabled the design of connected and integrated smart health monitoring systems. These health monitoring systems can be utilized for monitoring the mental and physical wellbeing of a person. Stress, anxiety, and hypertension are the major elements responsible for the plethora of physical and mental illnesses. In this context, the older population demands special attention because of the several age-related complications that exacerbate the effects of stress, anxiety, and hypertension. Monitoring stress, anxiety, and blood pressure regularly can prevent long-term damage by initiating necessary intervention or clinical treatment beforehand. This will improve the quality of life and reduce the burden on caregivers and the cost of healthcare. Therefore, this dissertation explores novel technological solutions for real-time monitoring of stress, anxiety, and blood pressure using unobtrusive wearable sensors and machine learning techniques.

The first contribution of this dissertation is the experimental data collection of 50 healthy older adults, based on which, the works on stress detection and anxiety detection have been developed. The data collection procedure lasted for more than a year. We have collected physiological signals, salivary

cortisol, and self-reported questionnaire feedback during the study. Salivary cortisol is an established clinical biomarker for physiological stress. Hence, a stress detection model that is trained to distinguish between the stressed and not-stressed states as indicated by the increase in cortisol level has the potential to facilitate clinical level diagnosis of stress from the comfort of their own home.

The second contribution of the dissertation is the development of a stress detection model based on fingertip sensors. We have extracted features from Electrodermal Activity (EDA) and Blood Volume Pulse (BVP) signals obtained from fingertip EDA and Photoplethysmogram (PPG) sensors to train machine learning algorithms for distinguishing between stressed and not-stressed states. We have evaluated the performance of four traditional machine learning algorithms and one deep-learning-based Long Short-Term Memory (LSTM) classifier. Results and analysis showed that the proposed LSTM classifier performed equally well as the traditional machine learning models.

The third contribution of the dissertation is to evaluate an integrated system of wrist-worn sensors for stress detection. We have evaluated four signal streams, EDA, BVP, Inter-Beat Interval (IBI), and Skin Temperature (ST) signals from EDA, PPG, and ST sensors. A random forest classifier was used for distinguishing between the stressed and not-stressed states. Results and analysis showed that incorporating features from different signals was able to reduce the misclassification rate of the classifier. Further, we have also

prototyped the integration of the proposed wristband-based stress detection system in a consumer end device with voice capabilities.

The fourth contribution of the dissertation is the design of an anxiety detection model that uses features from a single wearable sensor and a context feature to improve the performance of the classification model. Using a context feature instead of integrating other physiological features for improving the performance of the model can reduce the complexity and cost of the anxiety detection model. In our proposed work, we have used a simple experimental context feature to highlight the importance of context in the accurate detection of anxious states. Our results and analysis have shown that with the addition of the context-based feature, the classifier was able to reduce misclassification by increasing the confidence of the decision.

The final and the fifth contribution of the dissertation is the validation of a proposed computational framework for the blood pressure estimation model. The proposed framework uses features from the PPG signal to estimate the systolic and diastolic blood pressure values using advanced regression techniques.

KEYWORDS: Machine Learning, Physiological Signals, Stress, Anxiety, Blood Pressure, Salivary Cortisol.

Rajdeep Kumar Nath

July 20, 2021

NOVEL MACHINE LEARNING AND WEARABLE SENSOR BASED SOLUTIONS FOR SMART HEALTHCARE MONITORING

By

Rajdeep Kumar Nath

Dr. Himanshu Thapliyal
(Director of Dissertation)

Dr. Daniel Lau

(Director of Graduate Studies)

July 20, 2021

(Date)

ACKNOWLEDGEMENTS

This work was supported by the Kentucky Science and Engineering Foundation under Grant KSEF-3528-RDE-019.

Table of Contents

List of Figures

List of Tables

Chapter 1

Introduction

The general well-being of a person plays an important role in the daily life of an individual. Positive emotional well-being has been found to improve work performance, the quality of personal life, and develop emotional intelligence of a person [1][2][3]. However, daily stress and the associated anxiety can negatively impact the general well-being of a person both mentally and physically. Chronic stress is linked to the plethora of psychological and physiological illnesses such as anxiety disorder and hypertension [4]. This results in the increased cost of personal healthcare and also the burden on caregivers [5]. According to an estimate by the American Institute of Stress, job-related stress accounts for more than $ 300 billion per year [6].

According to a survey by American Psychological Association (APA), about 49% of the people reported that stress has negatively affected their behavior, and about 20% of the people reported that stress has caused ten-

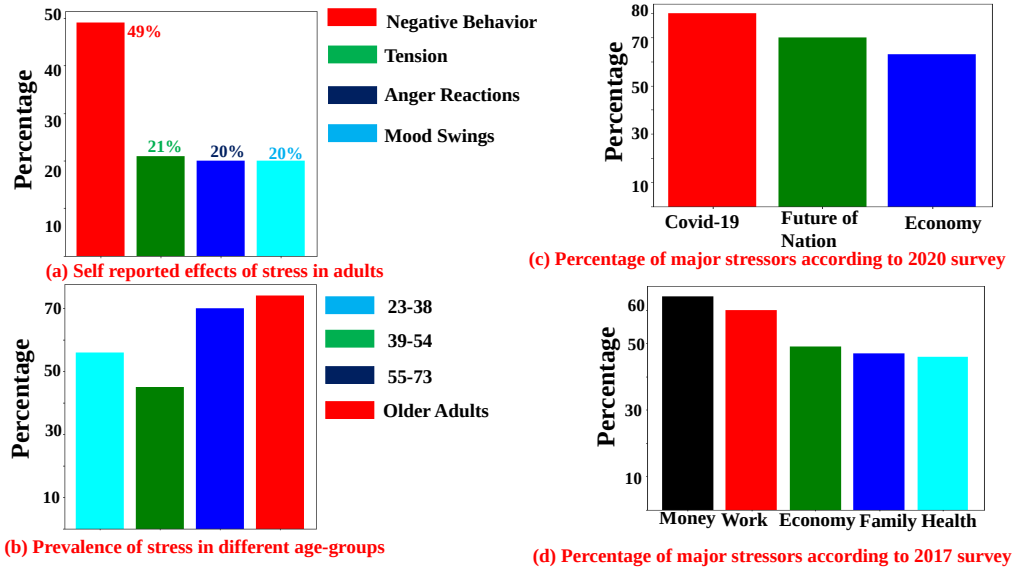
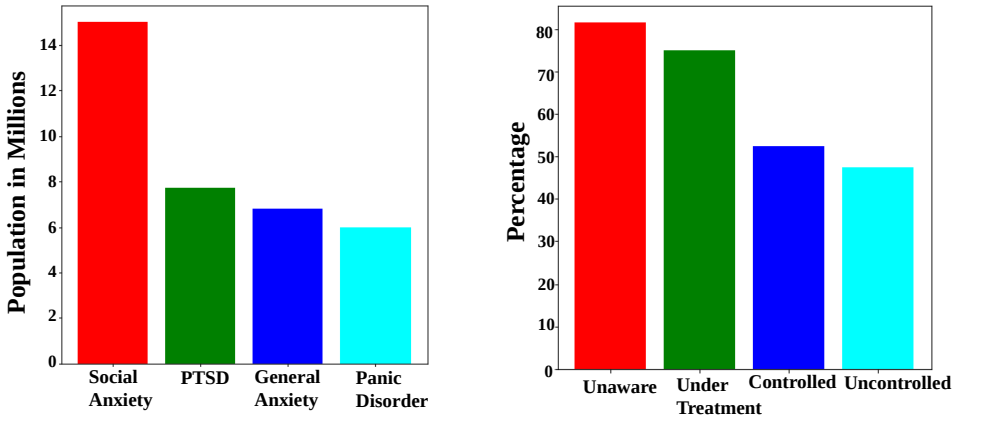


Figure 1.1: Various statistics on stress based on self-reported surveys

sion, mood swings, or anger reactions [7]. Stress is also prevalent among all age groups. As visualized from Figure 1.1 (b), it can be seen that older adults and adults above the age group 55 constitute the highest percentage of stressed individuals [7]. The major stressors responsible for stress in adults vary with the changing dynamics. For example, according to the 2020 survey, the emergence of Covid-19 was reported as the major source of stress (80% of the surveyed population), while according to the 2017 survey, money was reported as the major source of stress.

Apart from the direct social and economic effects of stress, chronic stress is also responsible for long-term anxiety and the development of high blood pressure over time [8][9]. Anxiety and high blood pressure also contribute negatively to the social and economic stability of the nation. Long-term



(a) Population (in millions) suffering from different kinds of anxiety disorder

(b) Percentage of people under different categories of high blood pressure

Figure 1.2: Statistics on anxiety and blood pressure based on surveys

exposure to anxiety causes anxiety disorder which can manifest itself in various forms (abnormalities). Figure 1.2(a) shows the population in millions suffering from different kinds of anxiety disorder such as social anxiety, post-traumatic stress disorder (PTSD), general anxiety, and panic disorder.

High blood pressure is another health emergency that is closely associated with anxiety and stress. The major concern with the development of high blood pressure is the lack of awareness of the condition, which further escalates to other severe health conditions. A survey conducted by the American heart association (AHA) found that about 80% of the people who have high blood pressure are unaware of the condition [7]. The same report states that about 47.5% of the people don't have their high blood pressure under control.

Stress, anxiety, and high blood pressure have a strong association with each other with some cause-and-effect relationship between them. In particular, older adults are disproportionately affected by stress. According to a report by American Psychological Association, older adults have experienced a greater increase in their stress levels as compared to other age groups [10]. Similarly, depression and generalized anxiety disorder are more prevalent among older adults [11][12]. According to a study conducted in [13], clinically significant symptoms of depression were found to be related to hypertension in older adults. In the next section, we will discuss these three elements in detail and discuss the association between them particularly in the context of older adults.

1.1 Stress, Anxiety, and Hypertension

Stress is defined as the reaction to adverse environmental situations that challenge the typical adaptive capability as perceived by an individual [14]. Although positive stress (eustress) helps the individual to stay focused to deal with adversities, negative stress (distress) causes the activation of the HPA (hypothalamic-pituitary-adrenocortical) axis. Repeated and prolonged activation of these pathways can increase the risk of premature aging, cognitive impairment, and cardiovascular and infectious diseases [15]

The adverse effect of stress in conjunction with the process of aging can result in irreparable damage to the brain and body over time [16]. Thus,

older adults are more susceptible to several stress-related ailments such as hypertension and anxiety disorder [17][4]. As in the aging process, stress can alter the functioning of the immune system [18]. Research suggests that both acute and chronic stress can lead to significant changes in the functionality of immune cells and this effect may worsen with age. Further, it has been found that an older, stressed individual is more prone to deficits in vaccine response, healing, recovery and may be more susceptible to infection-related illness as compared to a younger stressed individual [19][20]. Inadequate and improper management of natural episodic stress is found to adversely influence the immune function of older adults to a greater extent than that of relatively younger adults [21].

Anxiety is a complex emotional and behavioral response that results from the anticipation of a negative situation (real or perceived) that is potentially perceived as harmful to the individual [22][23]. Like stress, repeated exposure to anxiety can cause psychological disorders such as clinical depression and anxiety disorder [24]. Further, chronic anxiety has also been found to be strongly associated with physiological abnormalities such as cardiovascular diseases, insomnia, and impaired cognitive ability [25][26][27]. Although anxiety is a global psychological disorder affecting all population groups, a qualitative difference has been observed between the anxiety experienced by an older adult to that of a younger adult [28]. Statistics show that about 15% to 20% of older adults suffer from depression [29] and 15% to 40% of older adults suffer from anxiety [30]. The increased perception of anxiety and its

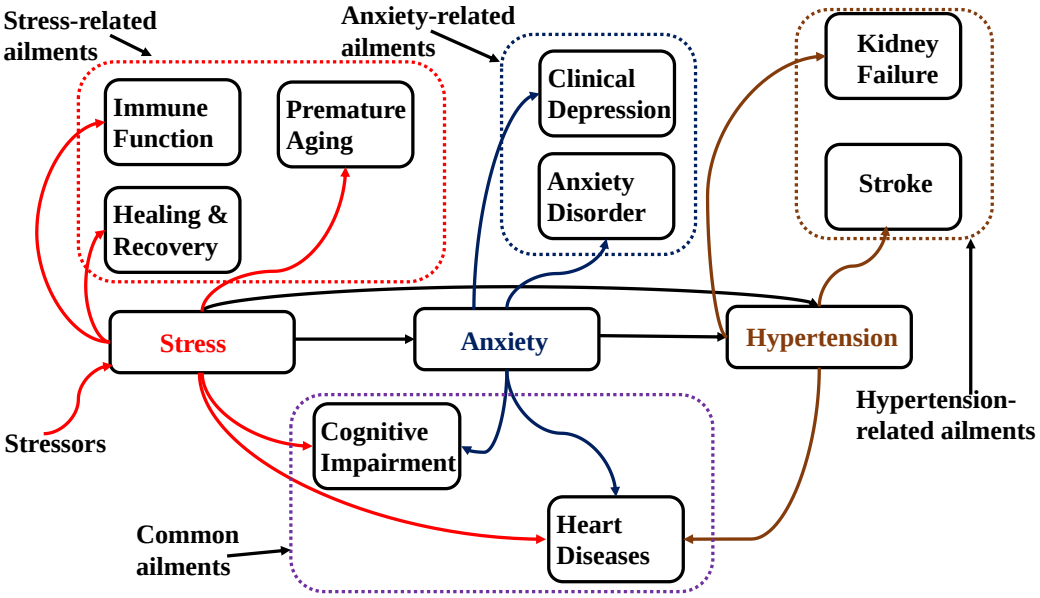


Figure 1.3: Concept map representing of the relationship between stress, anxiety, and hypertension and their related ailments

consequent effects in older adults is attributed to several factors prevalent among older adults such as loneliness from retirement, decreased physical abilities, etc. [31] [32].

Hypertension is another physiological condition that is irreversible. Abnormal or irregular blood pressure patterns over many years may lead to permanent conditions like hypertension. The development of hypertension in adults is a poorly understood topic because of several factors such as behavioral, genetic, and other psychosocial factors [33]. Hypertension is the leading cause and gateway to severe health problems like heart failure, vision loss, stroke, and complications related to kidney [34]. American Heart Association (AHA) estimates that about 85 million people (one in every 3 adults

over the age of 20) in the US have high blood pressure and according to the National Institutes of Health (NIH), about two-thirds of the people over the age of 65 have high blood pressure [34]. Studies have shown that there is some level of association between increased work-related load and high blood pressure [35]. Further, it was observed that improper management of work-life balance reflected by overburdening oneself with work can lead to a high level of blood pressure in everyday life which can eventually become chronic in nature [36][37].

From the above discussions on stress, anxiety, and hypertension, we can see that these three elements have overlapping health risks. In some cases, one health risk initiated by one element can lead to the development of another element and associated complicated factors. This relation between the three elements can be visualized with the help of the concept map in Figure 1.3. The visualization in Figure 1.3 represents some of the ailments of stress, anxiety, and hypertension. Stress has been found to play an important role in the development of both anxiety and hypertension [8][9]. Some studies have also found an association between anxiety and high blood pressure [38][39]. This concept is illustrated with arrow marks in Figure 1.3. Further, some stress-related ailments such as the development of cognitive impairment and heart diseases can also be caused by anxiety. Hypertension can also result in the development of heart diseases.

Hence, monitoring and managing stress and its associated effects such as anxiety and blood pressure can reduce the chances of the development

of chronic and irreversible health conditions later in life [40][41]. Moreover, logging stress, anxiety, and blood pressure profiles can help the caregiver in making an informed decision during diagnosis and early intervention can be initiated to prevent long-term damage. Further, the health data generated from these profiles can help the scientific community in better understanding the relationship between different ailments and the factors that lead to the development of such conditions. Moreover, these components can be integrated together with wearable location detection systems fall detection systems in a smart home environment context for providing additional context regarding the lifestyle of a person [42][43][44].

Advancement in low-power electronics and sensor technologies facilitates the integration of multiple wireless sensors in consumer wearable devices. Integration of consumer electronics component on the Internet of Medical Things (IoMT) for health monitoring using machine learning techniques is currently a widely researched topic [40][45][46][47][48]. Comfort is an important factor to be considered when designing health monitoring systems in a consumer electronics context [47]. Hence, a wearable system such as a wristband-based wearable system will be suitable for long-term use and is more likely to be used daily by the user [47].

In this dissertation, we will present novel machine learning and unobtrusive wearable sensor-based technological solutions to monitor stress, anxiety, and blood pressure in real-time. We will evaluate two wearable configurations for stress detection [49][50][51]. One of the wearable configurations is

a fingertip-based wearable system where the sensors are carefully attached to the user using velcro straps. Another configuration is a wristband-based wearable system. A wristband-based wearable system provides more freedom to the user in terms of usability. The anxiety detection model has been developed using the wristband-based wearable system [52]. Finally, a blood pressure estimation model using machine learning techniques is also proposed [53][54]. These technological solutions can facilitate real-time health monitoring and can provide scope for managing one's health through personalized intervention strategies and thereby take charge of one's well-being and reducing the need for regular intervention from a medical expert.

1.2 Contribution of Dissertation

The technical contributions of this dissertation are as follows:

1. **Experimental Data Collection:** Experimental data of 50 older adults were collected during the trier social stress test (TSST) experimental protocol for developing the stress and anxiety detection models. Physiological signals were collected using a fingertip-based wearable device and a smart wristband-based wearable device. Salivary cortisol, which is an established clinical biomarker for stress was collected for the ground truth estimation of stress levels. For anxiety level ground truth, the response from the STAI questionnaire was evaluated.

2. **Stress Detection Model Based on Fingertip-Based wearable**

Device: A stress detection model using fingertip EDA and PPG sensors and traditional machine learning algorithms along with a proposed long short-term memory (LSTM) is proposed for distinguishing between stressed and not-stressed states. Experimental data of 19 older adults were used for this work.

3. **Stress Detection Model Based on Smart Wristband-Based Wearable Device:** A stress detection model using wristband EDA, PPG, and ST sensors was used along with a random forest classifier machine learning model to distinguish between the stressed state and not-stressed state. Experimental data of 40 older adults were used for this work. Further, the proposed stress detection model is prototyped in a consumer end device with voice capabilities, so that users can receive feedback on their vitals and stress levels by querying on voice-enabled consumer devices such as smartphones and smart speakers.

4. **Anxiety Detection Using Wearable Sensor and Context Feature:** An anxiety detection model that uses a single wearable sensor (EDA and PPG) along with an experimental context feature is proposed. Experimental data of 41 older adults were used for this work.

5. **Blood Pressure Estimation Using Wearable Device:** A blood pressure estimation and classification model using a PPG sensor is proposed. Data used for building the blood pressure estimation model is extracted from an online freely available MIMIC dataset.

1.3 Outline of Dissertation

This dissertation is structured into six main chapters excluding the current chapter that is the introduction to the dissertation. Chapter 2 will discuss the physiological signals, wearable sensors, and machine learning techniques for building the predictive models, the performance metrics used for evaluating the predictive models, and the related work in stress detection, anxiety detection, and blood pressure estimation. Chapter 3 will present the technical work on developing stress detection models based on fingertip sensors such as EDA and PPG. Chapter 4 will present the technical work on developing stress detection models based on smart wristband-based sensors such as EDA, PPG, and ST. Chapter 5 will present the work on anxiety detection using a single physiological sensor and context feature. Chapter 6 will present the work on blood pressure estimation and classification. Finally, Chapter 7 will conclude the dissertation.

Chapter 2

Background

In this chapter, we will present the background of this dissertation. This chapter is organized as follows: Section 2.1 will discuss the fundamentals of physiological signals for health monitoring in the context of stress detection, anxiety detection, and blood pressure estimation. Section 2.2 will discuss the wearable devices for health monitoring in the context of stress detection, anxiety detection, and blood pressure estimation. Section 2.3 will discuss the fundamentals of machine learning techniques used for predictive modeling in the context of stress detection, anxiety detection, and blood pressure estimation. Section 2.4 will discuss the performance metrics used for quantifying the performance of the predictive models. Section 2.5 will discuss some of the representative works on stress, anxiety, and blood pressure estimation. Finally, section 2.6 will conclude the chapter.

2.1 Physiological Signals for Health Monitoring

In this section, we will discuss the various physiological measures that can be used for the continuous detection of stress, anxiety, and blood pressure. Continuous recording of physiological signals using wearable sensors for monitoring stress, anxiety, and blood pressure is getting attention from research community [55][56][57]. Common physiological signals that are used in the context of stress, anxiety, and blood pressure are galvanic skin response (GSR) also known as electrodermal activity (EDA), photoplethysmogram (PPG) also known as Blood Volume Pulse (BVP), electrocardiogram (ECG), electroencephalogram (EEG), skin temperature (ST), and Respiration (RESP).

Galvanic skin response (GSR) or electrodermal activity (EDA) is the most common physiological measure for stress detection). GSR is related to physiological and psychological arousal. Arousal in the autonomic nervous system (ANS) increases the activity of the sweat glands resulting in increased skin conductance. Figure 2.1 shows how GSR is related to the activation of the ANS to aroused state and deactivation from the stressed state to relaxed state [1]. Usually, for the best quality GSR signal, the sensors are placed on the fingertip or foot. However, measurement using a wrist-based GSR sensor is more practical for ambulatory purposes [58].

Photoplethysmogram (PPG) is also known as blood volume pulse (BVP) in the literature. PPG signal provides information regarding cardiac activ-

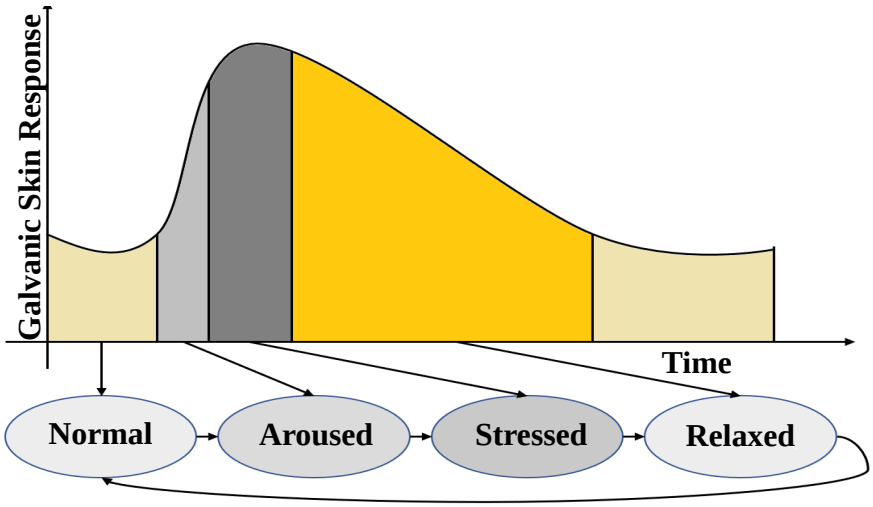


Figure 2.1: Variation of GSR with mental stress [1] (© 2020 IEEE)

ity. PPG signal is a periodic signal with subsequent systolic and diastolic peaks. PPG signal has been used for stress, anxiety, and blood pressure estimation along with combination of other signals ([59][55][51]). Electrocardiogram (ECG) is another periodic signal that is a measure of cardiac activity. ECG signal has also been used for detecting stress and anxiety and also for estimating blood pressure values in the literature [57][55][60].

Electroencephalogram (EEG) captures the brain activity signal and has the potential to reflect the emotional arousal during stressful situations [61]. Because of this reason, the EEG signal is a popular choice for detecting anxiety and stress [59][62]. However, capturing EEG signals requires a complex acquisition setup which might not be comfortable for the user in an ambulatory setting. Skin temperature is also another useful signal that is associated with stress and anxiety [63]. The skin temperature signal is also

a convenient choice because it can be captured relatively easily either by adopting fingertip-based devices or by adopting wristband-type devices that use thermal imaging for recording the skin temperature without remaining in direct contact with the skin [64][65]. Respiration is also another useful physiological signal that can detect arousal during stressful situations, which is usually indicated by the increase in the respiration rate. Respiration signal is usually captured using breath sensors. However, the respiration signal is relatively complicated to capture as any movement or talking associated during recording can reduce the quality of signal [66].

2.2 Wearable Sensors for Health Monitoring

Several prototypes and implementation methods of wearable frameworks for recording multiple physiological signal monitoring have been proposed in the literature. In the work by Yoon et al.,[67] researchers have developed a monitoring patch that is capable of capturing skin temperature, skin conductance, and pulse wave signals. In the work by Lee et al.,[68] a glovebased sensor to capture EDA signals and pulse wave signals has been proposed. In the work by Fletcher et al.,[69] a physiological signal monitoring and communication system were developed by researchers from MIT Media Lab. A smart sensor capable of capturing heart rate, skin conductance, and skin temperature was proposed by Quazi et al.,[70]. In the work by Healey and Picard [71] a stress monitoring system based on a body sensor network has been designed for an

ambulatory setting. However, there are also several commercially available devices and platforms for physiological signal acquisition and recording. Table 2.1 lists some popular commercially available devices suitable for data collection and analysis for research in the area of stress detection.

Table 2.1: Popular commercially available devices for research (© 2020 IEEE).

Brand	**Device**	**Signals**	**RTI**	**Ambulant**
Empatica	E4 wristband	PPG, GSR, HR, ACC, ST	Yes	Yes
Garmin	Vivosmart	HR, HRV, ACC	Yes	Yes
Zephyr	BioHarness 3.0	HR, HRV, GSR, ACC, ST	Yes	Yes
iMotions	Shimmer 3+ GSR	GSR, PPG	Yes	No
BIOPAC	Mobita Wearable	ECG, EEG, EGG EMG, and EOG	Yes	No

GSR = Galvanic Skin Response, HR = Heart Rate, ACC = Acceleration, ST = Skin Temperature, HRV = Heart Rate Variability, PPG = Photoplethysmograph, RTI = Real Time Implementation

In the next section we will discuss how machine learning techniques are used in conjunction with signal processing techniques to quantify stress, anxiety, and blood pressure levels.

2.3 Machine Learning Techniques for Health Monitoring

Features from physiological signals represent the relationship between these physiological signals and corresponding ground truth labels. Ground truth labels represent the actual value of the independent variable. For example, stress detection and anxiety detection are typically formulated as binary classification problems, where these binary values (such as stressed or not-stressed, anxious or not-anxious) are the ground truth labels. A machine learning algorithm maps the features from physiological signals to the ground truth labels and builds a classification model based on the training data. The problem of blood pressure estimation is typically formulated as a regression problem, where the ground truth labels are the minimum and maximum arterial blood pressure (ABP) values that represent the systolic and diastolic blood pressure values respectively.

Figure 2.2 shows the general machine-learning-based framework adopted in the literature for building predictive models for stress, anxiety, and blood pressure. The three main stages for building a machine-learning-based predictive model from physiological signals are preprocessing, feature extraction, and training and testing. We will discuss these stages briefly in this section.

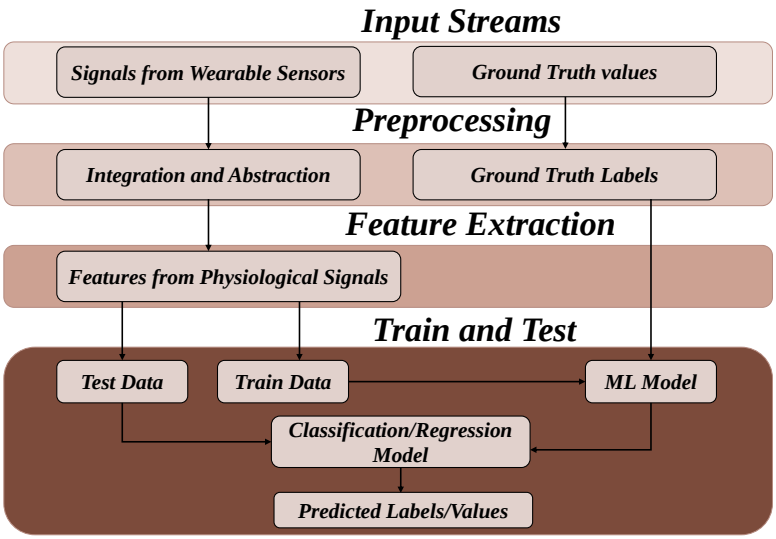


Figure 2.2: Overview of a machine learning based framework for stress, anxiety, and blood pressure estimation (© 2020 IEEE).

2.3.1 Preprocessing

In this stage, the raw signals from the wearable sensors are processed to remove noise and other outliers. Common computational steps in this stage consist of normalization of the raw signals into some specific range. Then usually a low-pass filter is applied to remove the high-frequency components from the signal. The processed signal is then used for extracting features from the physiological signals.

2.3.2 Feature Extraction

The process of feature extraction is analogous to mapping the high-dimensional raw physiological signals into some low-dimensional feature vector. Features

are usually computed by using a rolling windowing technique usually with overlapping time segments. Researchers have used a combination of these signals for developing predictive models for detecting stress, anxiety, and blood pressure. It has been observed that the sensor fusion technique (concept of fusing information from different sensors) to develop predictive models usually outperforms the predictive models that use a single sensor [72][73]. Features extracted from physiological signals vary significantly based on applications and also the signals used. Researchers have used a combination of features from different signal streams such as GSR, ECG, EEG, and RESP. Some of the common features extracted from these signals for stress detection are represented as a taxonomical diagram in Figure 2.3

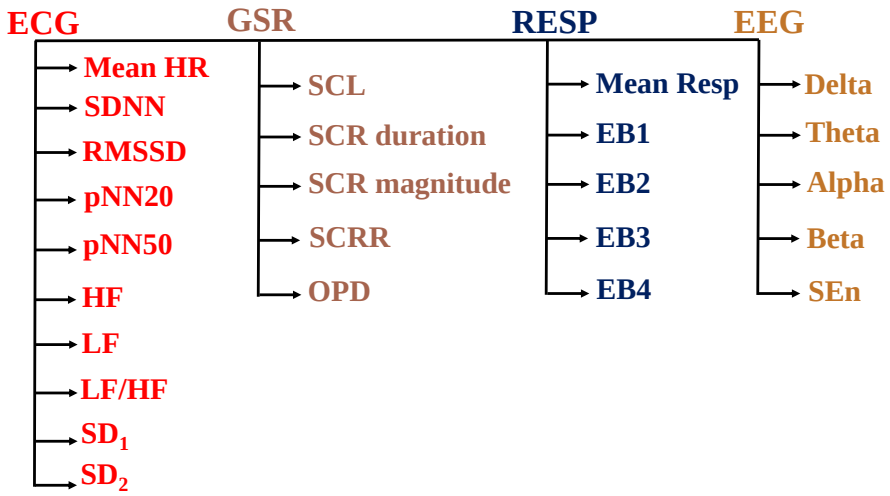


Figure 2.3: Common features extracted from ECG, GSR, Respiration and EEG for stress detection (© 2020 IEEE).

For anxiety detection, EEG and ECG based features are more studied

in the literature [46][74][59]. However, features from signals such as PPG, eye movement, audio signals, and tri-axial acceleration have also been used in the context of anxiety detection. For blood pressure estimation, however, combinations of mostly cardiovascular signals such as PPG, ECG, phonocardiogram (PCG), and ballistocardiography (BCG). For blood pressure estimation, features such as pulse transit time (PTT), pulse arrival time (PAT), and pulse wave velocity (PWV) are computed. Along with traditional approaches, using deep learning techniques for feature extraction and prediction has also received attention in the research community [75].

2.3.3 Training and Testing

For building the predictive model, the features are trained using a machine learning algorithm. For problems such as stress detection and anxiety detection, a classification algorithm is adopted and for problems such as blood pressure estimation, a regression algorithm is adopted. The algorithm classifies or categorizes a feature sample to a target class or target value depending on the feature values and the classification rule. Table 2.2 shows the characteristic of some popular machine learning algorithms typically used in the context of stress detection, anxiety detection, and blood pressure estimation.

Table 2.2 shows the different attributes of machine learning algorithms such as the decision rule used by the classifier, the training, and the testing/prediction complexity. It is interesting to note that the prediction or testing complexity of these classifiers depends on the number of features

Table 2.2: Characteristics of popular machine learning algorithms (© 2020 IEEE).

Algorithm	**Classification Rule**	**Training**	**Testing**
SVM	Support Vectors	$O(n^2p + n^3)$	$O(n_{SV}p)$
κ-NN	Distance Criteria	NA	$O(np)$
DT	Decision Tree	$O(n^2p)$	$O(p)$
LDA	Dimension Reduction	NA	$O(npt + t^3)$
NB	Bayes Theorem	$O(np)$	$O(p)$

n = number of samples, p = number of features, n_{SV}=number of support vectors, $t = min(n, p)$,SVM=Support Vector Machine,κ-NN=κ Nearest Neighbor, DT=Decision Tree, LDA=Linear Discriminant Analysis, NB=Naive Bayes

(p). Hence, selecting the most accurately representing features through feature selection criteria such as Pearson correlation, mutual information score between the feature variable and target variable is a standard practice to reduce the complexity of the learning algorithm and also to improve the performance of the predictive model. Recently, quantum annealing was also used for feature selection for stress detection as well as for some other application domains [76][77]. In [78], the problem of stress detection has been modeled for detecting future stress events rather than the current stress events.

Having discussed the fundamentals of building a predictive model, we will now present the discussion on evaluating the performance of the predictive models.

2.4 Performance Metrics

In this section, we will discuss the performance metrics used for quantifying the performance of machine learning models. For classification tasks, the performance metrics mostly used are the accuracy, F1-scores, and AUC score. In the case of regression tasks such as blood pressure estimation, the performance metrics used are the mean absolute error (MAE) and standard error. We will describe these performance metrics separately in the context of classification and regression.

2.4.1 Classification

During binary classification tasks, the objective of the machine learning classifier is to assign a class label to the feature sample. Out of the two classes, one class is taken as the positive class and the other negative. For example, in the case of stress detection, the stressed ('S') class can be thought of as the positive class and the not-stressed (NS) class can be thought of as the negative class. Similarly for anxiety detection, anxious ('A') is the positive class and not-anxious (NA) is the negative class. Hence, during one prediction task, there can be four possible outcomes, true positive (TP), true negative (TN), false positive (FP), and false negative (FN).

In our results and analysis, the performance of the machine learning model is quantified using three metrics: (i) macro-average f1-score, (ii) micro-average f1-score, and (iii) Area under curve (AUC) score. If TP_p, FP_p and FN_p represents the true positive, false positive and false negative of the positive class('S' or 'A') and TP_n, FP_n and FN_n denotes the same for the negative class ('NS' or 'NA'), then, micro-average precision is given by the equation.

$$P_{micro} = \frac{TP_p + TP_n}{TP_p + TP_n + FP_p + FP_n} \tag{2.1}$$

and micro-average recall is given by the equation:

$$R_{micro} = \frac{TP_p + TP_n}{TP_p + TP_n + FN_p + FN_n} \tag{2.2}$$

The resulting micro-average f1-score can be calculated as

$$F_{micro} = \frac{2.(P_{micro}.R_{micro})}{(P_{micro} + R_{micro})} \tag{2.3}$$

Macro-average precision is given by:

$$P_{macro} = \frac{2.TP_p.TP_n + TP_p.FP_n + TP_n.FP_p}{2.(TP_p + FP_p).(TP_n + FP_n)} \tag{2.4}$$

Macro-average recall is given by:

$$R_{macro} = \frac{2.TP_p.TP_n + TP_p.FN_n + TP_n.FN_p}{2.(TP_p + FN_p).(TP_n + FN_n)} \tag{2.5}$$

Hence, macro-average f1-score can be calculated by:

$$F_{macro} = \frac{2.(P_{macro}.R_{macro})}{(P_{macro} + R_{macro})} \tag{2.6}$$

Macro average weighs both the positive class and the negative class equally while micro average favors the majority class [79]. AUC score represents the area under the ROC (receiver operating characteristic) curve and is given by the equation:

$$AUC = \frac{1}{2}.(\frac{TP}{TP + FN} + \frac{TN}{TN + FP}) \tag{2.7}$$

where TP, TN, FP, FN represents true positive, true negative, false positive and false negative of the confusion matrix.

2.4.2 Regression

For regression tasks such as blood pressure estimation, the machine learning model assigns a discrete value to the feature sample. The accuracy is calculated in terms of Mean Absolute Error (MAE) which quantifies how much the predicted value deviates from the actual value on an average. Another metric commonly used for regression tasks is the Standard Error (SE) which quantifies the incorrectness of the regression model on an average as it provides an estimate of how far apart are the predicted values on an average from the regression line. In other words, it is the standard deviation of the errors observed on a set of measurements.

Mathematically MAE and SE are defined as:

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|x_i - x| \tag{2.8}$$

$$SE = \sqrt{\frac{\sum(x_i - \bar{x})^2}{n-1}} \tag{2.9}$$

, where x_i is the predicted value and x is the observed value and n is the number of observation.

2.5 Related Work on Stress, Anxiety, and Blood Pressure Estimation

Having discussed the fundamentals of detecting stress, anxiety, and blood pressure, we will now explore the related works in these areas. As these works have been conducted under different experimental conditions, we will perform a qualitative evaluation of the related works.

2.5.1 Stress Detection

Detecting stress levels from physiological signals has been gaining immense importance in the scientific community in recent years. However, most of the existing works have either used younger adults or no such target group has been specifically mentioned. Another area that has been relatively less explored is the use of cortisol as the stress reference for stress detection. Most of the work has used stress levels quantified from self-reporting questionnaires like DASS 21 (Depression, Anxiety and Stress Scale - 21), etc., or from other subjective methods as the ground truth. The response obtained by such a method may be biased due to inaccurate perceptions of stress, low levels of objectivity, and inconsistencies in assessment [80][81]. Work done in [56] has used cortisol as the ground reference for stress detection. However, the target population was younger adults and the study involved a small participant pool. Comfort is also another important factor to be considered when designing a stress monitoring system in a consumer electronics context [82].

For example, wrist-based acquisition devices are likely to be more comfortable for long-term monitoring especially for older adults. A study conducted in [82] showed that older adults are more likely to show a positive attitude in terms of comfort in using a wrist-worn device than younger adults and continue to use and adapt to the device. Our work provides a novel direction for detecting stress in older adults using cortisol as the stress reference and using fingertip and wrist-worn sensors to monitor physiological signals.

Table 2.3: Attributes of related work on stress detection in older adults.

Work	**Subjects**	**Acquisition Points**	**Signals**	**Target Population**	**Stress Reference**
Kikhia et al. [83]	6	Wrist	EDA	Older Adults With Dementia	Annotation by Clinical Staff
Belk et al. [84]	Use Case	Hand/Palm	EDA, IMU, ST HR, Grip Force	Older Adults	Annotation by External Observer
Delmastro et al. [60]	9	Chest, Finger	ECG, EDA	MCI Frail Older Adults	Predetermined Stress Annotation
Cheong et al. [85]	9	Wrist	ST, Step Count, HR, Humidity Near Body and Air Temperature	Older Adults	Self Reported Questionnaire
Ferreira et al. [62]	17	Finger, Arm Chest, Head	ECG, EEG EDA	Older Adults	Self Reported Questionnaire
Our Work. [49]	**19**	**Fingertip**	**EDA, BVP**	**Older Adults**	**Salivary Cortisol**
Our Work. [51]	**40**	**Wrist**	**EDA, BVP, IBI, ST**	**Older Adults**	**Salivary Cortisol**

Table 2.3 lists some of the related work and their key attributes in the area of stress monitoring targeting older adults. In [83], researchers have used a wrist-based EDA sensor to detect stress levels in older adults with dementia. The study was conducted on 6 participants and the stress reference was obtained from the annotations provided by clinical staff. In [84],

EDA, IMU, HR, and Grip Force were used to detect stress levels in older adults. A computer mouse equipped with the corresponding sensor was used to record the signals. In [60], ECG signal recorded from chest and EDA signal recorded from finger was used to detect stress. The study was performed on 9 participants who were suffering from Mild Cognitive Impairment. In [85], a combination of skin, near-body, and air temperature along with step count and relative humidity data were used to detect stress. In [62], researchers have used data from Finger, Arm, Chest, and Head to record ECG, EEG, and EDA signals. The study was conducted on 17 older adults and the ground truth was accessed by a self-reported questionnaire. Table 2.3 also contains the qualitative attributes of our fingertip and wristband-based work on stress detection for comparison.

2.5.2 Anxiety Detection

Technological solutions to detect anxiety in real-time using physiological signals have recently garnered significant attention from the research community. Unlike stress detection, anxiety detection is relatively less explored in the context of the Internet of Medical Things (IoMT). For example, there is a significant amount of research work available in the literature that uses low-cost consumer electronic components to monitor stress continuously in real-time [86][51]. However, there are very few research articles that detail technological solutions for detecting anxiety using low-cost consumer electronic components. This work explores a novel direction for anxiety detection

using wrist-worn sensors and a context-aware feature. Our work explores an interesting solution towards realizing a low-cost consumer electronic system to monitor anxiety in real-time for older adults.

Table 2.4: Attributes of related work on Anxiety detection in older adults.

Work	**Subjects**	**Acquisition Points**	**Signals**	**Target Population**	**Anxiety Reference**
Puli et al. [74]	15	Chest and arm	ECG, ACC	Children and Youth with ASD	Physiological Indicators
Wen et al. [57]	59	Wrist and ankle	ECG	Younger Adults	Annotation by Audience Score
Zhang et al. [59]	92	Head	EEG, eye-movement	Children and Adolescents	Self-report (SCARED)
Li et al. [46]	12	Head	EEG	Younger Adults	Predetermined levels
Zheng et al. [87]	20	Head and nose	EEG, PPG	Younger Adults	STAI
McGinnis et al. [88]	71	NA	Audio	Children	Self-reported and observed
This Work [52]	**41**	**Wrist**	**Wrist-based signals (EDA and BVP)**	**Older Adults**	**STAI**

Table 2.4 presents some of the existing works in anxiety detection along with a qualitative comparison with our proposed work. In [74], Puli et al. proposed a Kalman-like filter approach that fuses heart rate features extracted from ECG signal along with accelerometer signals for accurate anxiety level detection in the presence of motion. The target population for this work was children and youth who are clinically diagnosed with ASD (Autism Spectrum Disorder). The experiment was performed with 15 participants and the results and the analysis from the experiment showed that fusing accelerometer data with heart rate features was able to reduce false detection of anxiety to a significant level. Wen et al. [57] proposed an anxiety detection model that uses features from the IBI series extracted from the ECG signal. The anxiety detection model proposed by the researchers

uses a trained SVM model to classify samples as high or low anxiety states based on the extracted ECG features. In [59], Zhang et al. proposed fusing eye-movement features with EEG signals to increase the precision of anxiety detection. The researchers used group sparse canonical correlation analysis (GSCCA) to generate effective feature space containing both EEG and eye-movement features for improved performance of the anxiety detection model. The features were trained using SVM. Results and analysis showed that the fusing features from eye movement and EEG were able to improve the performance of the anxiety detection model.

In [46], Li et al. have used extensive feature set from the EEG signal such as frequency domain, time domain, non-linear, and statistical features to train an SVM model to distinguish anxiety in 4 levels. The experiment was performed on 12 university students who were younger adults. A sensor fusion approach to classify anxiety in three levels was explored by Zheng et al. [87] using EEG and PPG signals. 20 younger adults were used in this experiment during which the participants were required to complete two cycling-related tasks. Results from their analysis showed that features from both EEG and PPG performed better in detecting anxiety than when features from a single signal was used. In [88], McGinnis et al. explored the effectiveness of audio data in detecting anxiety among young children. Audio features were extracted from the audio signal recorded during a 3-minute speech task and the performance of several machine learning algorithms such as logistic regression, random forest, and SVM were evaluated.

2.5.3 Blood Pressure Estimation

Majority of the research on estimating blood pressure using physiological signals focussed on the analysis of Pulse Transit Time (PTT), Pulse Arrival Time (PAT), and Pulse Wave Velocity (PWV). However, to calculate these quantities, either a multiple sensing modality or dual element probe is required. For example, in [89], researchers have used a dual element PPG probe to capture PPG data which in turn calculates Pulse Wave Velocity. Work done in [55] used pulse transit time and pulse arrival time to estimate blood pressure values by measuring three signals, Electrocardiogram (ECG), Phonocardiogram (PCG), and Photoplethysmogram (PPG). In [90], the authors have used Ballistocardiogram (BCG) for measuring pulse transit time which is used for estimating blood pressure values. In our work, we have explored a novel technique to estimate blood pressure values by using a single sensor and single probe PPG-based computational framework.

2.6 Conclusion

In this chapter, we have discussed some of the common physiological signals in the context of stress, anxiety, and blood pressure estimation. We have also discussed the fundamentals of machine learning techniques commonly used for building predictive models for stress, anxiety, and blood pressure. We have presented the discussion on performance metrics used for evaluating the machine-learning-based predictive models for stress, anxiety, and blood

pressure. Finally, the related work in the area of stress detection, anxiety detection, and blood pressure estimation has been discussed.

Chapter 3

Stress Detection Using Fingertip Sensors

In this chapter we will present the work on stress detection using fingertip sensors. In this work, we will evaluate the effectiveness of a fingertip-based wearable physiological stress monitoring system in distinguishing between stressed and non-stressed state in older adults using machine learning techniques. This system utilizes EDA and BVP signal to detect occurrence of stress as indicated by salivary cortisol measurement which is a reliable objective measure of physiological stress. Data of 19 healthy older adults (11 female and 8 male) with mean age 73.15 $\pm$ 5.79 were used for this study. EDA and BVP signals were recorded using a finger tip sensor during the trier social stress test (TSST), which is a well known experimental protocol to reliably induce stress in humans in a social setting. Four machine learning

algorithms were evaluated based on their performance in classifying between stressed and non-stressed states. The four machine learning models are Random Forest (RF), κ-Nearest Neighbor (κ-NN), Logistic Regression (LR), and Support Vector Machine (SVM). Further, the effectiveness of a novel deep learning Long Short-Term Memory (LSTM) based classifier in distinguishing between stressed and non-stressed state.

In Section 3.1, we will discuss the experimental design adopted for the study. This section will discuss the selection criteria used for screening participants for the study, the methods used for recording and collecting physiological data and the experimental protocol. In Section 3.2, we will discuss the signal processing and features extraction procedure used to extract features from EDA and BVP signal. The processing of salivary cortisol samples for stress annotation will be discussed in Section 3.3. In Section 3.4, we will present the supervised feature selection algorithm and the machine learning models used for correlating the features with stress classes. Section 3.5 will present the results and analysis. Section 3.6 presents some discussion on the proposed work and finally Section 3.7 concludes the chapter.

3.1 Experimental Study Design

The objective of this experiment is to correlate the features from EDA and BVP signals with increase in cortisol concentration in older adults when exposed to stressors. In this section, we will discuss the design of our experi-

mental study. We will start our discussion with the screening criteria adopted for recruiting participants for the study. This will be followed by a discussion on the materials and methods used for physiological data collection. Finally, the experimental protocol adopted for this study will be discussed.

3.1.1 Participant Screening

Healthy older adults between the ages of 60 to 80 were recruited for this study. Participants were excluded from participation in the study if they reported having existing heart conditions, post traumatic disorder, anxiety disorders, unstable angina, Addison's disease, or Cushing disease. 19 participants, 11 female and 8 male with mean age 73.15 ± 5.79 were enrolled in this study.

3.1.2 Physiological Data Collection

Physiological signals and salivary cortisol samples were collected during the study. The physiological signals captured for this study were the EDA and BVP signals. EDA also known as Galvanic Skin Response (GSR) measures the electrical conductivity of the skin. This is measured by applying a low and constant voltage using two electrodes to skin.

The second physiological signal considered for our study is BVP also known as Photoplethysmogram (PPG). BVP contains information about the cardiac activity. Although Heart Rate (HR) and Heart Rate Variability (HRV) features are usually extracted from ECG signal for better accuracy,

these features can also be approximated using BVP signal as Pulse Rate Variability (PRV) is considered to be highly correlated to HRV [91]. Further, EDA and BVP are very convenient and cost-effective to measure and require minimum obtrusiveness.

To record EDA and BVP signal, Shimmer3 GSR+ has been used. The EDA sensors are electrodes of roughly 1 cm^2 attached to a Velcro strap secured on the middle and ring finger. The two sensors are used as a signal/source pair for measuring conductivity (responsiveness to electricity). The BVP sensor configuration is composed of a small (roughly $1cm^2$) clear epoxy covered circuit that has a light source and detector with a Velcro strap to which the BVP sensor is fixed. The sensor and strap are secured firmly to the tip of the index finger. These electrodes are placed on the non dominant hand of the participant to minimize signal distortion during the experiment. The data captured by the sensors is transmitted in real time via a Bluetooth connection to a dedicated software platform for storing the data for future analysis. The signals are analyzed and processed on a computer with Intel i5 processor and 16 GB RAM.

Saliva samples are taken five times during the study at time stamps T1, T2, T3, T4, and T5 (Figure 3.1). T1 is taken when the participant just arrives at the study facility. T2 is taken at the end of Pre Stress period. T3 is taken after the stress induction period, T4 is taken after the first recovery phase and T5 at the end of the study. Cortisol response has a lag of 20 minutes [92], hence the samples are taken at an interval of 20 minutes. Salivary samples

are taken using a cotton swab which the participants are instructed to put the swab under the tongue for about 2 minutes. The swabs are then stored at about -20 deg Celsius. The participants were instructed not to eat or drink anything before the study as it could affect the quality of the cortisol concentration.

3.1.3 Experimental Protocol

For the experimental protocol, the Trier Social Stress Test (TSST) has been selected because of its capability to capture social and psychological stress in a naturalistic environment [93, 94]. Figure 3.1 illustrates the steps in the experimental protocol. The protocol consists of a waiting period and a Pre-Stress (PS) period for 20 minutes (time period T1-T2), Anticipatory Stress (AS) period of 10 minutes followed by Speech and Math task for another 10 minutes (time period T2-T3). During the waiting period, participants have been asked to complete demographic surveys, and clarify any questions about the study process they may have. The physiological signal recording starts from the PS period. During the PS period, the baseline measurement of the participant is taken which is for approximately 2 minutes. After that, the participants are shown a topic on which they have to speak continuously for 5 minutes while being observed. This is the AS period which lasts for 10 minutes. During this time, the participant is expected to mentally prepare the speech. After the AS period, there are subsequent stress phases: (i) The speech task and (ii) mental math task, each of which lasts for 5 minutes each.

During the mental math task, participants are asked to solve simple addition and subtraction problems mentally while being observed. The math task is set up such that the participants must answer each question correctly before proceeding to the next question. The difficulty level of the math equations increases with every correct answer. The final steps in the TSST protocol are two recovery periods which are 20 minutes each. (time period T3-T5).

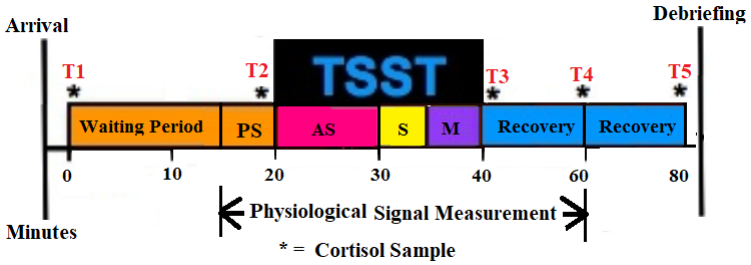


Figure 3.1: Experimental Protocol (© 2021 Springer Nature Switzerland AG)

3.2 Signal Processing and Feature Extraction

In this section, we will discuss the methods used for processing physiological signal and feature extraction. The overview of the complete framework used for stress classification is shown in Figure 3.2.

Features extracted from EDA and BVP signals form the backbone of the stress detection model on which the machine learning model trains itself to classify between stressed and not stressed states. Before extracting features, the raw EDA signal in uS (Microsiemens) and BVP signal in mV (Millivolts) is scaled between 0 and 1 and subsequently a low pass Butterworth filter of

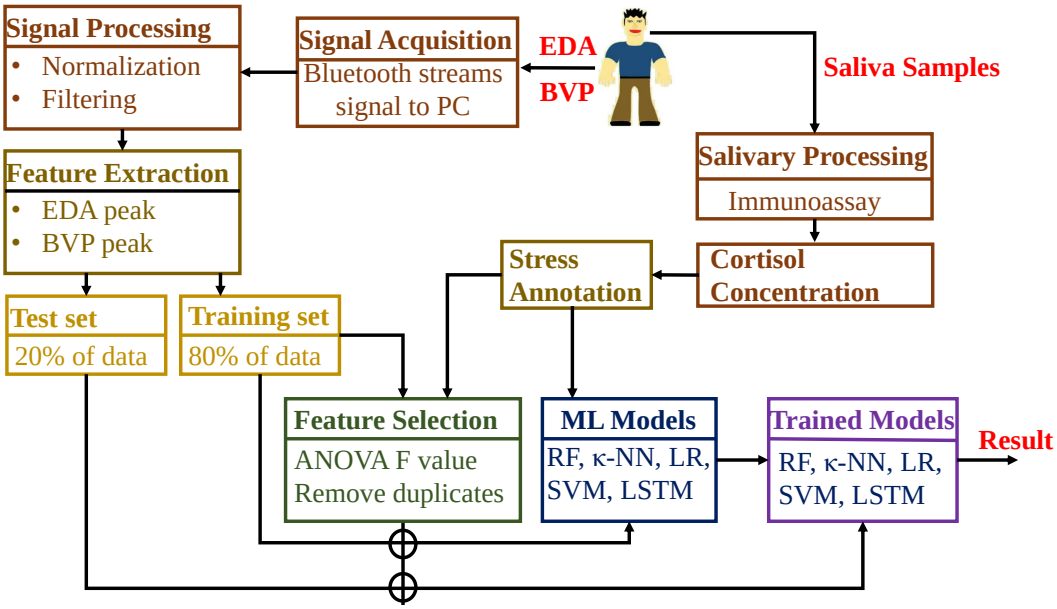


Figure 3.2: Overview of the stress classification framework used in this work. ML models represent the models prior to weight estimation. The ML models are trained using the stress annotation from cortisol concentration and the training feature set. The trained models are then tested using the test feature set (© 2021 Springer Nature Switzerland AG).

order 5 and cutoff frequency 1 Hz for EDA signal and 10 Hz for BVP signal is used.

Table 3.1: Description of EDA extracted features (© 2021 Springer Nature Switzerland AG)

Symbol	Description
P_{amp}^{MEAN}	Mean of peak amplitudes
P_{amp}^{MEDIAN}	Median of peak amplitudes
P_{amp}^{SD}	Standard deviation of peak amplitudes
P_{amp}^{RMS}	Root mean square of peak amplitudes
P_{amp}^{MAX}	Maximum of peak amplitudes
P_{amp}^{MIN}	Minimum of peak amplitudes
P_{width}^{MEAN}	Mean of peak widths
P_{width}^{MEDIAN}	Median of peak widths
P_{width}^{SD}	Standard deviation of peak widths
P_{width}^{RMS}	Root mean square of peak widths
P_{width}^{MAX}	Maximum of peak widths
P_{width}^{MIN}	Minimum of peak widths
P_{prom}^{MEAN}	Mean of peak prominence
P_{prom}^{MEDIAN}	Median of peak prominence
P_{prom}^{SD}	Standard deviation of peak prominence
P_{prom}^{RMS}	Root mean square of peak prominence
P_{prom}^{MAX}	Maximum of peak prominence
P_{prom}^{MIN}	Minimum of peak prominence

3.2.1 EDA Feature Extraction

EDA is characterized by occurrence of peaks also known as SCR (skin conductance response) peaks. Research evidence suggests that skin conductance response is associated with emotional arousal [95, 96]. Research conducted in [73] showed that the peak amplitude and instantaneous peak rate contains information about the stress level of the person. Based on these research evidence we hypothesize that observing the peak characteristic during stressed and non stressed condition can help contribute in training a machine learning model in classifying between the stressed and non-stressed state. We have attempted to study three characteristic of SCR peak: (i) peak amplitude, (ii) peak width and (iii) peak prominence. Prominence of a peak is a quantitative measure of how much a peak is distinguished from its surrounding baseline and is measured by the vertical line between the peak and the lowest contour line. The observation of these peak characteristic was summarized by computing statistical measures over the 90 sec running window and overlap of 45 sec.

The peak detection algorithm [97] is applied to the first derivative of EDA signal and the detected peaks are visualized with varying threshold. A threshold of 0.001 seemed appropriate in accurately detecting most of the peaks. Statistical measures such as mean, median, standard deviation, root mean square, maximum, and minimum of peak amplitude, peak width and the prominence of the peak were computed as features. Table 3.1 describes the extracted features from EDA signal. The features extracted from EDA

signal is represented by the symbol $P^{statistics}_{quantity}$, where *quantity* represents what aspect of the peak we are measuring like amplitude, width and prominence and *statistics* represents what statistical measure is being computed. A total of 18 statistical features are computed from EDA signal.

3.2.2 BVP Feature Extraction

BVP signal is characterized by two types of peaks: (i) Systolic and (ii) Diastolic peaks. A single cardiac cycle is composed of both these peaks. During stressful situations, cardiovascular arousal can result in changes in heart rate, blood pressure etc [98]. Work done in [99] showed that the systolic and diastolic time changes under stressful situation. Similarly stress situations can influence the morphology of systolic and diastolic peaks. Based on these evidence, we have computed features based on the characterization of these peaks in terms of amplitude, width, prominence and time difference between consecutive peaks. In this work, we have only considered the characterization of systolic peaks. This is because Systolic peaks can be more accurately detected than diastolic peaks because of the relatively high amplitude of systolic peak and hence it can be easily separated from the peaks arising out of motion artifacts.

Peak detection algorithm was applied to processed BVP signal using a 90 second running window and overlap of 45 second to detect systolic peaks. The detected peaks were visualized under varying thresholds and a threshold value of 0.4 seemed appropriate in accurately detecting the systolic peaks.

Furthermore, the distance and width between two consecutive peaks is set to $f_s * 0.65$ and $f_s * 0.1$ respectively, where f_s is the sampling frequency. This is because a typical systolic phase lasts for 0.3 second and diastolic phase lasts for about 0.4 second [100]. Based on this, a rough assumption about the threshold for the peak width and distance can be estimated. As period of systole lasts for about 0.3 seconds, it can be assumed that the width of a peak to be classified as peak should be atleast half the width of the entire systole. Hence, the threshold for the width is set to 0.1 sec. Further, as the time difference between two systoles is about 0.7 sec, the time difference between two consecutive peaks should be atleast around 0.65. Hence, the threshold for peak distance is selected to be about 0.65 seconds.

Statistical measures such as mean, median, standard deviation, root mean square, maximum, and minimum of peak width and the prominence of the peak were computed as features. Also, statistical measure such as mean, standard deviation and root mean square of the time difference between two consecutive peaks and the number of peaks appearing per minute were calculated. Finally mean, standard deviation, root mean square and range of peak amplitude is computed to form the final BVP feature set. The features extracted from BVP signal is represented by the symbol $S^{statistics}_{quantity}$, where *quantity* represents what aspect of the peak we are measuring like amplitude, width, prominence, and time interval and *statistics* represents what statistical measure is being computed. A total of 21 statistical features are computed from BVP signal. Table 3.2 shows the description of the extracted

features from BVP signal.

Table 3.2: Description of BVP extracted features (© 2021 Springer Nature Switzerland AG).

Symbol	Description
S_{min}	No. of peaks per minute
S_{time}^{SD}	Standard deviation of the time differences between two consecutive peaks
S_{time}^{RMS}	Root mean square of the time differences between two consecutive peaks
S_{time}^{Mean}	Mean of the time differences between two consecutive peaks
S_{time}^{Range}	Difference between the maximum and minimum time differences between two consecutive peaks
S_{width}^{MEAN}	Mean of peak widths
S_{width}^{MEDIAN}	Median of peak widths
S_{width}^{SD}	Standard deviation of peak widths
S_{width}^{RMS}	Root mean square of peak widths
S_{width}^{MAX}	Maximum of peak widths
S_{width}^{MIN}	Minimum of peak widths
S_{prom}^{MEAN}	Mean of peak prominence
S_{prom}^{MEDIAN}	Median of peak prominence
S_{prom}^{SD}	Standard deviation of peak prominence
S_{prom}^{RMS}	Root mean square of peak prominence
S_{prom}^{MAX}	Maximum of peak prominence
S_{prom}^{MIN}	Minimum of peak prominence
S_{amp}^{MEAN}	Mean of peak amplitudes
S_{amp}^{SD}	Standard deviation of peak amplitudes
S_{amp}^{RMS}	Root mean square of peak amplitudes
S_{amp}^{Range}	Range of peak amplitudes

3.3 Salivary Cortisol Processing and Stress Annotation

The objective of salivary cortisol processing is to convert the continuous cortisol values to categorical values, stressed (S) or not-stressed (NS). The salivary samples were processed and cortisol concentration were estimated. The assay range was 0.012-3 (ug/dL). For each participant, there were 5 cortisol samples, two of which were taken during pre- study period (T1 and T2) and three of which were taken during the study (T3-T5). The following preprocessing steps were implemented to classify a sample as either stressed or not-stressed.

- Samples taken at T1 and T2 were averaged and then subtracted from the remaining samples at taken T3, T4 and T5 for each participant.
- The three samples taken at T3, T4 and T5 for each participant were integrated and standardized by subtracting the mean from each sample and dividing by the standard deviation.
- A sample x was classified as stress (S) if $x > \mu$ and as not stressed (NS) if $x < \mu$, where μ is the population mean.

The statistics for stress class (S) is 0.31 ± 0.27 (ug/dL) and that of not-stressed class (NS) is -1.17 ± 1.64 (ug/dL) with p-value of $5.33e^{-7}$. p-value is used to quantify the difference between the two groups. Usually a p-value less

than 0.05 suggests strong evidence that the null hypothesis can be rejected and the alternate hypothesis can be accepted. In this case as the p-value is $5.33e^{-7}$, which is significantly less than 0.05, we can accept the alternative hypothesis that the stressed and non-stressed groups are significantly different from each other.

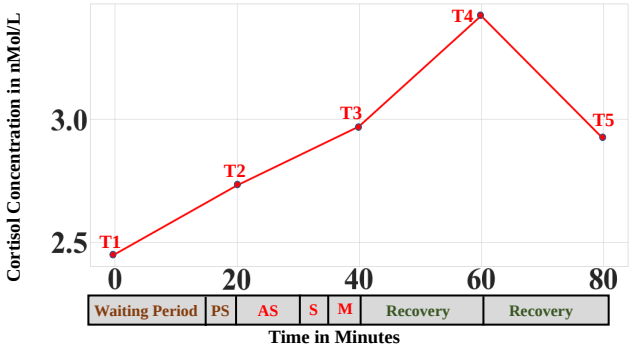


Figure 3.3: Example of the trend of cortisol concentration during the study (© 2021 Springer Nature Switzerland AG).

Figure 3.3 shows an example of the trend of cortisol concentration during the study session. We can see from Figure 3.3 that the cortisol concentration followed an increasing trend upon participant exposure to stressors and the relative concentration decreased during in the recovery phase. The mean and the standard deviation of the cortisol concentration for each of the five time stamps for the 19 participants is shown in Table 3.3.

Table 3.3: Mean and SD (Standard Deviation) of cortisol concentration during each time stamps (© 2021 Springer Nature Switzerland AG).

Time Stamps	Mean (ug/dL)	SD (ug/dL)
T1	0.107	0.059
T2	0.149	0.145
T3	0.142	0.121
T4	0.131	0.092
T5	0.117	0.079

3.4 Feature Selection and Machine Learning Models

In this section, we will discuss the process of feature extraction and present discussion on the machine learning models which will be used to correlate the EDA and BVP features to stress classified by salivary cortisol processing.

3.4.1 Feature Selection

Table 3.4: Rank of the features selected along with the signal source (© 2021 Springer Nature Switzerland AG).

Rank	Feature	Source	Rank	Feature	Source	Rank	Feature	Source
1	S_{amp}^{Range}	BVP	11	P_{prom}^{MIN}	EDA	21	S_{prom}^{MAX}	BVP
2	S_{prom}^{MEDIAN}	BVP	12	P_{width}^{MIN}	EDA	22	S_{amp}^{MEAN}	BVP
3	S_{prom}^{MIN}	BVP	12	P_{prom}^{MEAN}	EDA	23	P_{width}^{MEDIAN}	EDA
4	S_{width}^{SD}	BVP	14	P_{width}^{MEAN}	EDA	24	P_{width}^{SD}	EDA
5	S_{amp}^{SD}	BVP	15	P_{prom}^{MAX}	EDA	25	S_{width}^{MIN}	BVP
6	S_{prom}^{MEAN}	BVP	16	P_{amp}^{MAX}	EDA	26	P_{width}^{MEAN}	EDA
7	S_{width}^{MAX}	BVP	17	P_{prom}^{SD}	EDA	27	S_{time}^{MEAN}	BVP
8	S_{prom}^{SD}	BVP	18	P_{width}^{MAX}	EDA	28	P_{amp}^{SD}	EDA
9	S_{min}	BVP	19	P_{prom}^{MEDIAN}	EDA	29	S_{time}^{RMS}	BVP
10	P_{width}^{MIN}	EDA	20	S_{prom}^{MIN}	BVP	30	P_{amp}^{MEAN}	EDA

The feature set is split into train and test set in a 80-20 ratio while ensuring that no two samples from the same participants are both in the training and test set.

The feature selection algorithm (Algorithm 1) takes input the train data and outputs the index of the selected features. The steps involved in the feature selection algorithm is explained with line numbers. Line 1 and 2 denotes the assignment of training feature vector and training target vector to X_train and Y_train respectively. In line 3, score for each feature is calculated using the ANOVA F [97] value between the feature and the target class. In line 4, the score is sorted and the duplicate scores are removed in

Algorithm 1: Feature Selection Algorithm (© 2021 Springer Nature Switzerland AG).

Input : Train Data
Output: Index of selected features

```
1 X_train ← training feature vector ;
2 Y_train ← training target data;
3 score ← f_classif(X_train, Y_train);
4 sorted_score ← sort(score);
5 sorted_score ← drop_duplicates(score);
6 for i in (0, len(sorted_score) do
7 |   index_select_feature ← indexof(sorted_score[i]);
8 end
9 return index_select_feature
```

line 5. Then from line 6 to line 8, index of selected features are extracted and returned in line 9.

The feature selection algorithm returned 30 features out of the 39 features extracted from EDA and BVP signal. Table 3.4 shows the rank of the features according to their score and their source. The final feature set contains 15 features each from EDA and BVP.

3.4.2 Machine Learning Models

For correlating physiological features with stress, four types of machine learning algorithms are studied: (i) Decision tree based classifier, (ii) Nearest neighbor based classifier, (iii) Probability based classifier and (iv) Kernel based classifier. Random Forest (RF) classifier is chosen as the decision tree based classifier, κ Nearest Neighbor (κ-NN) is chosen as nearest neighbor

based classifier, Logistic Regression (LR) is chosen as the probability based classifier and finally Support Vector Machine (SVM) is chosen as the kernel based classifier.

Apart from these four standard popular machine learning classifiers, we have also evaluated the performance of a deep learning based classifier. Long Short-Term Memory (LSTM) is chosen as the deep learning based classifier. The following section discusses the architecture of the proposed LSTM network.

3.4.3 Proposed LSTM Network

In this section, we will present and discuss the architecture of the proposed LSTM network.

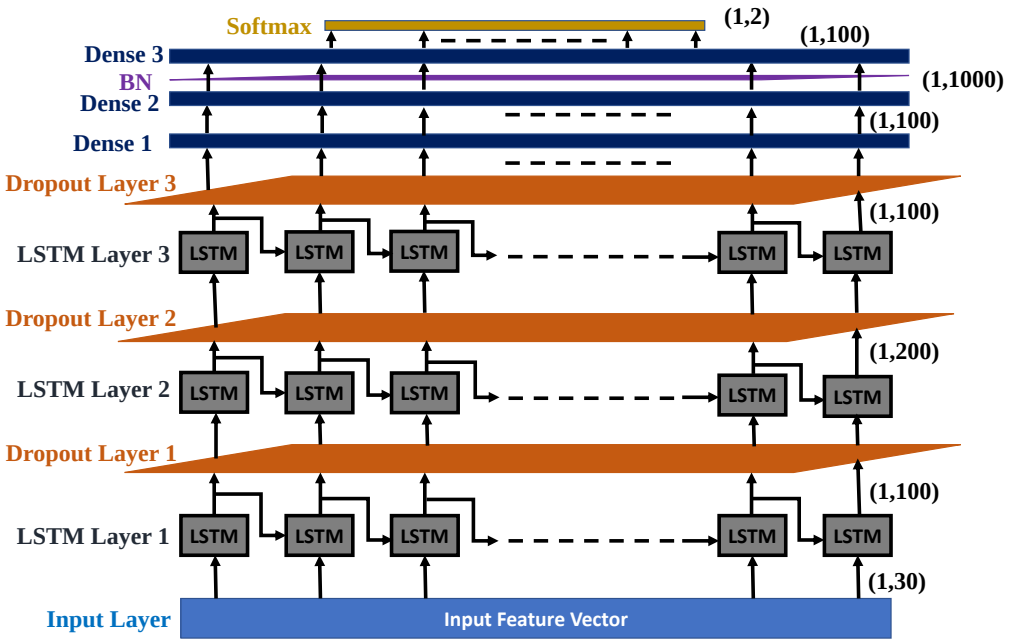


Figure 3.4: Schematic Layout of the Proposed LSTM Architecture. The numbers in bracket represent the dimension of the vector output in each layer. (BN=Batch Normalization) (© 2021 Springer Nature Switzerland AG).

The proposed LSTM architecture consists of three stacked LSTM layers followed by three Dense layers. Figure 3.4 shows the layout of different layers in the network. The first layer is an LSTM layer with 100 hidden units. The input shape for each sample is $(1, 30)$ where 30 is the number of selected feature and 1 is the time step. The output from the first LSTM layer is activated by sigmoid function. The next layer is the dropout layer. In this layer, 20% of the neurons chosen at random are dropped before propagating to the next stage. The purpose of the dropout layer is to prevent over fitting. The third layer is again an LSTM layer with 200 hidden units and activated

by sigmoid function. This layer is again followed by a dropout layer which again connects to the last LSTM layer with again 100 hidden units followed by another dropout layer. The output from the last dropout layer is forwarded to the first fully connected (dense) layer of the network with 1000 neurons. The output of this dense layer is again sigmoid activated before forwarding to the Batch Normalization layer. The output from the BN layer then is forwarded to another dense layer with 100 neurons. the output from this layer is again activated by sigmoid function and forwarded to the last dense layer for classification activated by softmax function.

3.5 Results

In this section, we will discuss the results of the machine learning models in classifying between stressed and not-stressed states.

3.5.1 Performance of RF, κ-NN, LR and SVM

To estimate the model parameters and tuning hyperparameters, k fold cross validation technique is applied with k=5 to randomly generate training and validation dataset. A 5 fold cross validation was deemed suitable for our analysis, proportional to our training set with each folds having sufficient number of training samples. Finally the trained model was tested on the test data. Table 3.5 shows the performance of the four ML models.

Table 3.5: Performance evaluation of ML classifiers (© 2021 Springer Nature Switzerland AG).

ML Models	Macro Precision	Macro Recall	Macro F1-score	Micro F1-score	AUC Score
RF	0.94	0.53	0.53	0.88	0.53
LR	0.98	0.81	0.87	0.95	0.81
SVM	0.75	0.59	0.61	0.86	0.60
κ-NN	0.95	0.56	0.58	0.88	0.56

It can be seen from Table 3.5 that logistic regression model performed the best among the other ML algorithms achieving a macro-average f1-score of 0.87, micro-average f1-score of 0.95 and an AUC score of 0.81. The micro-average f1-score for all the four models are significantly high, whereas, macro-average score of RF, κ-NN, and SVM is 0.52, 0.57 and 0.61 respectively which is significantly low when compared to that of LR model which is 0.87. This shows that RF, κ-NN and SVM are heavily influenced by class imbalance and hence is more biased towards the majority class. Further, the AUC score for LR model is found to be 0.81 which is also significantly greater than those of RF, κ-NN and SVM which is 0.53, 0.56 and 0.60 respectively. Hence, based on AUC score, we can say that LR classification model is better at handling false classification than RF, κ-NN or SVM.

3.5.2 Performance of LSTM Network

For training the LSTM model, the training feature set is further randomly divided into 80-20 ratio with 20% of the data reserved for validation and 80%

reserved for training. To avoid over fitting and optimize training time, the robustness of the model is validated after every epoch. This is done by implementing two call back criteria: (i) Early Stopping and (ii) Model Checkpoint. The early stopping criteria monitors the validation loss and keeps track of the minimum loss. If the loss does not improve after training for 250 epochs, the callbacks criteria forces the training to stop. The model checkpoint criteria is used to monitor the validation accuracy for its maximum value and updates its model weights whenever the validation accuracy increases and the model weights are saved. The LSTM model has been trained for 2000 epochs with 256 samples as the batch size and with categorical cross entropy as the loss function and adam as the optimizer. Epochs refer to the number of times the training sample is passed through the neural network. Figure 3.5 shows the training and validation loss and accuracy with number of epochs.

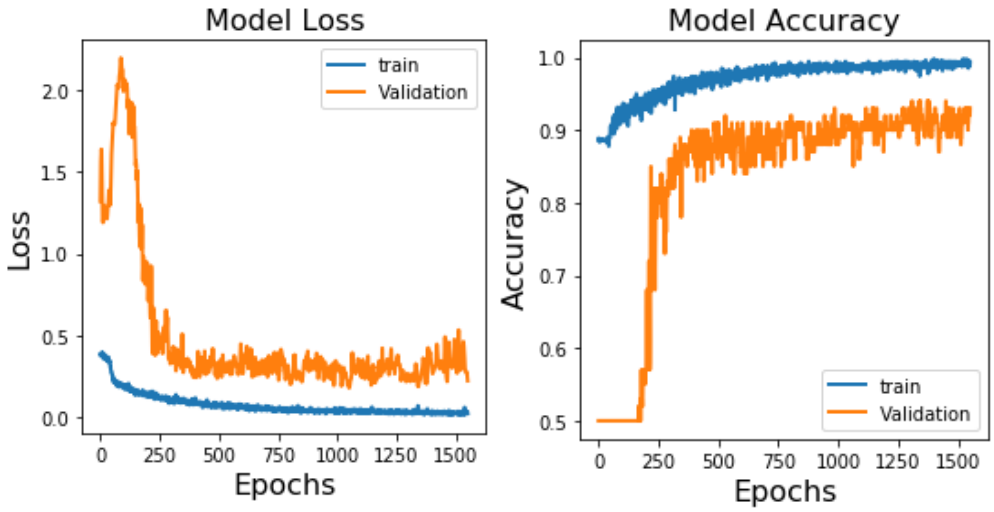


Figure 3.5: Plot of the training and validation loss and accuracy with the number of epochs (© 2021 Springer Nature Switzerland AG).

We can see from Figure 3.5 that the validation loss and accuracy starts to converge with the training loss and accuracy after about 250 epochs. The validation loss stopped improving after 1250 epochs and the training was stopped at 1500 epochs.

The weights of the best saved model is used to evaluate the performance of the LSTM network. The LSTM model achieved a macro average f1-score of 0.93, micro average f1-score of 0.97 and an AUC score of 0.90. When compared to the performance of the logistic regression model, we can see that LSTM classifier improved the macro average f1-score by 6.7%, micro average f1-score by 2% and AUC score by 11%. The overall accuracy achieved is 97%. Table 3.6 shows the precision, recall and f1-score for both the classes.

Results show that LSTM based classifiers could also be a potential choice to be used for stress level classification alongside standard machine learning models.

Table 3.6: Performance measure on test data (© 2021 Springer Nature Switzerland AG).

Class	**Precision**	**Recall**	**F1-score**
NS	0.93	0.81	0.87
S	0.98	0.99	0.98
	Macro F1-score	**Micro F1-score**	**AUC Score**
	0.93	0.97	0.90

3.6 Discussion

The objective of this study is to evaluate the effectiveness of detecting objective stress levels as indicated by increase in salivary cortisol concentration using only EDA and BVP signals. These signals can be recorded easily by wrist-worn devices or fingertip sensors which makes real time monitoring of stress convenient and unobtrusive. To the best of our knowledge this is the only work that has evaluated the effectiveness of the features of EDA and BVP alone to detect objective stress as defined by cortisol concentration. Work done in [56] also used cortisol as the ground truth, however the use of EEG signal in detecting stress may not be suitable and convenient to monitor stress continuously especially for older adults as recording EEG requires obtrusive setup of an EEG headset which may lead to the added frustration and anxiety of older people.

3.7 Conclusion

In this chapter, we have evaluated the effectiveness of an unobtrusive fingertip-based wearable device system for physiological stress detection in older adults. This system utilizes EDA and BVP signals to extract statistical measures of the peak characteristic of EDA and BVP signal. The TSST Protocol has been used to induce stress in 19 participants including 11 females and 8 males and with mean age 73.15 ± 5.79 The features are trained and tested with four machine learning algorithms. Results indicated that logistic regression performed the best achieving macro-average f1-score, micro-average f1-score and AUC score of 0.87, 0.95 and 0.81 respectively. Further a deep learning based LSTM classifier is proposed and evaluated on the test set. Results showed that LSTM based classifier could improve the micro-average f1-score and macro-average f1-score by 2% and 6.7% respectively from the best performing logistic regression classifier. Further, LSTM classifier obtained an AUC score of 0.90 which is about 11% higher than the AUC score of logistic regression. This work illustrates the potential of the proposed system in building a physiological stress detection device using EDA and BVP signal to unobtrusively and continuously monitor stress level of an individual. Such a device could be specifically useful to closely monitor the stress levels of older adults in the comfort of their homes, without having to rely on clinical tests, thereby facilitating the capability of real time feedback of stress levels. This work also will also allow for the study of personalized

intervention techniques to counteract the harmful effects of stress on health. Ultimately, by helping to reduce stress this work may aid in the prevention of cognitive decline in older adults, thereby, reducing the risk of Mild Cognitive Impairment, Alzheimer's Disease, and other stress-related illnesses.

Chapter 4

Stress Detection Using Smart Wristband

In the previous chapter, we presented our work on detecting stress using fingertip-based EDA and PPG sensors. Fingertip-based EDA and PPG sensors might not be best choice for real time monitoring in an ambulatory setting because of the associated comfort and poor signal quality due to motion artifact [58]. However, a system of wristband-based wearable sensors is expected to be more suitable for continuous monitoring in an ambulatory setting [58][32]. In this chapter we will explore a novel approach for detecting stress using wristband-based sensors.

In this work, our objective is to design, develop, and evaluate the effectiveness of a stress detection model for older adults using a system of wrist-worn sensors. Our system uses four signals, EDA, BVP, IBI, and ST

from EDA, PPG, and ST sensors, embedded in a smart wristband, to classify between stressed and not-stressed state. The stress reference is obtained from salivary cortisol measurement, which is a well established clinical biomarker for measuring physiological stress. This work is the result of year-long data collection and analysis of 40 older adults (28 females and 12 males) and age 73.625 ± 5.39. EDA, BVP, IBI, and ST signals were collected during TSST (Trier Social Stress Test), which is a well known experimental protocol to reliably induce stress in humans in a social setting. Further, we prototype the proposed stress detection model in a consumer end device with voice capabilities, so that users can receive feedback on their vitals and stress levels by querying on voice-enabled consumer devices such as smartphones and smart speakers.

This chapter is organized as follows, section 4.1 presents the proposed method for stress detection. In this section, we will present the method of feature extraction from physiological signals, processing of salivary cortisol, calibration of salivary cortisol concentration with physiological signals, and the training and testing procedure. Section 4.2 presents the discussion on the experimental set up for the study. Section 4.3 presents the results and analysis of the study. Section 4.4 discusses the integration of the proposed stress detection system in a consumer electronics framework, and finally Section 4.5 concludes the chapter.

4.1 Proposed Method for Stress Detection

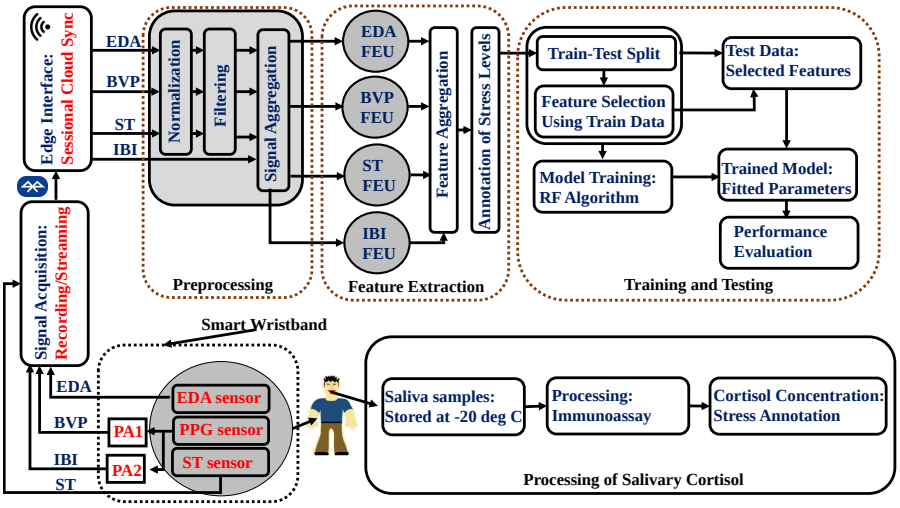


Figure 4.1: Overview of the Proposed Method used for Stress Detection. EDA signal is obtained from EDA sensor, BVP, and IBI are obtained using two proprietary algorithms in the device, and ST sensor captures the skin temperature. FEU=Feature Extraction Unit, PA=Proprietary Algorithm (© 2021 IEEE)

The overview of the proposed method for detecting stress from four physiological signals: (i) EDA (Electrodermal Activity), (ii) BVP (Blood Volume Pulse), (iii) IBI (Inter-Beat Interval), and (iv) ST (Skin Temperature) can be visually represented in Figure 4.1. The main components of the proposed method are (i) Preprocessing, (ii) Feature Extraction, (iii) Processing of Salivary Cortisol (iv) Training and Testing. We will discuss each of the following components in this section.

4.1.1 Preprocessing

The raw EDA and BVP signals are first normalized so that the values are in the range [0,1]. Subsequently, a low pass Butterworth filter of order 5 is used for filtering the EDA and BVP signals. Signal components higher than 1 Hz were removed for EDA signals, and signal components higher than 10 Hz were removed for BVP signals. ST signal was also first normalized in the range [0,1] and filtered to remove outliers or replace missing values etc. No preprocessing on the IBI signal is performed as it is already a smoothed signal obtained from the PPG sensor by the acquisition device.

4.1.2 Feature Extraction

Features are extracted from EDA, BVP, IBI, and ST signals after the preprocessing stage. The features from all the signal streams are extracted using a running window of 90 sec and an overlap of 45 sec. The notation of the extracted feature and their description is presented in Table 4.1

EDA Feature Extraction Unit

This unit extracts features from the EDA signal. EDA signal captures the electrodermal activity, which is a measure of the electrical variation of the skin as a result of arousal. This variation is modulated by the activity of the sweat gland which is under the control of the sympathetic nervous system [101]. Characterization of EDA peaks or startles has been shown to contain

information regarding the stress level of a person [73]. Peak detection algorithm [97] was applied on the first derivative of the EDA signal and the peak amplitude, peak width, and peak prominence were computed. These peak characteristics are summarized by computing descriptive statistics of the peak amplitude, width, and prominence. A total of 18 features were extracted from the EDA signal. The symbol of EDA features are represented by $P^{statistics}_{quantity}$, where *quantity* represents the quantitative measure of the peak that is being computed such as amplitude, width, or prominence and *statistics* represents the statistical measure that is being computed.

BVP Feature Extraction Unit

This unit extracts features from the BVP signal. BVP signal measures cardiovascular activity. The BVP signal is obtained from the PPG sensor. BVP signal is associated with cardiovascular arousal which is responsible for changes in heart rate and blood pressure during stressful situations [98]. Peak detection algorithm [97] has been used to detect systolic peaks. A total of 17 features were extracted from the BVP signal. Similar to EDA, statistical summary of peak amplitude, width, and prominence were computed as BVP features. Along with these, the number of systolic peaks occurring per minute is also computed. The symbol of the features are represented as $S^{statistics}_{quantity}$

IBI Feature Extraction Unit

This unit extracts features from the IBI signal. IBI or Inter Beat Interval is the time difference between two consecutive beats of the heart. A total of 6 statistical features were extracted from the IBI signal all in the time domain.

ST Feature Extraction Unit

This unit extracts features from the ST signal. A total of 6 features are computed from the ST signal. Out of the six features, four features are statistical measures of the skin temperature and the remaining two features are the slope and intercept of a fitted regression line.

Table 4.1: Notation and Description of the extracted features from EDA, BVP, IBI and ST (© 2021 IEEE).

Signal Source	Feature Notation	Feature Description	Signal Source	Feature Notation	Feature Description
EDA	$P_{amp}^{\bar{x}}$	Mean of peak amplitudes	BVP	$S_{min}^{\#x}$	No. of peaks per minute
EDA	$P_{amp}^{\tilde{x}}$	Median of peak amplitudes	BVP	$S_{width}^{\bar{x}}$	Mean of peak widths
EDA	P_{amp}^{σ}	Standard deviation of peak amplitudes	BVP	$S_{width}^{\tilde{x}}$	Median of peak widths
EDA	$P_{amp}^{\sqrt{x^2}}$	Root mean square of peak amplitudes	BVP	$S_{width}^{\sigma x}$	Standard deviation of peak widths
EDA	$P_{amp}^{\vee x}$	Maximum of peak amplitudes	BVP	$S_{width}^{\sqrt{x^2}}$	Root mean square of peak widths
EDA	$P_{amp}^{\wedge x}$	Minimum of peak amplitudes	BVP	$S_{width}^{\vee x}$	Maximum of peak widths
EDA	$P_{width}^{\bar{x}}$	Mean of peak widths	BVP	$S_{width}^{\wedge x}$	Minimum of peak widths
EDA	$P_{width}^{\tilde{x}}$	Median of peak widths	BVP	$S_{prom}^{\bar{x}}$	Mean of peak prominence
EDA	P_{width}^{σ}	Standard deviation of peak widths	BVP	$S_{prom}^{\tilde{x}}$	Median of peak prominence
EDA	$P_{width}^{\sqrt{x^2}}$	Root mean square of peak widths	BVP	$S_{prom}^{\sigma x}$	Standard deviation of peak prominence
EDA	$P_{width}^{\vee x}$	Maximum of peak widths	BVP	$S_{prom}^{\sqrt{x^2}}$	Root mean square of peak prominence
EDA	$P_{width}^{\wedge x}$	Minimum of peak widths	BVP	$S_{prom}^{\vee x}$	Maximum of peak prominence
EDA	$P_{prom}^{\bar{x}}$	Mean of peak prominence	BVP	$S_{prom}^{\wedge x}$	Minimum of peak prominence
EDA	$P_{prom}^{\tilde{x}}$	Median of peak prominence	BVP	$S_{amp}^{\bar{x}}$	Mean of peak amplitudes
EDA	P_{prom}^{σ}	Standard deviation of peak prominence	BVP	$S_{amp}^{\sigma x}$	Standard deviation of peak amplitudes
EDA	$P_{prom}^{\sqrt{x^2}}$	Root mean square of peak prominence	BVP	$S_{amp}^{\sqrt{x^2}}$	Root mean square of peak amplitudes
EDA	$P_{prom}^{\vee x}$	Maximum of peak prominence	BVP	$S_{amp}^{\vee x \Delta \wedge x}$	Range of peak amplitudes
EDA	$P_{prom}^{\wedge x}$	Minimum of peak prominence	IBI	$IBI^{\bar{x}}$	Mean of IBI
IBI	$IBI^{\tilde{x}}$	Median of IBI	ST	$ST^{\bar{x}}$	Mean of ST
IBI	$IBI^{\sigma x}$	Standard deviation of IBI	ST	$ST^{\tilde{x}}$	Median of ST
IBI	$IBI^{\sqrt{x^2}}$	Root mean square of IBI	ST	$ST^{\sigma x}$	Standard deviation of ST
IBI	$IBI^{\vee x}$	Maximum of IBI	ST	$ST^{\sqrt{x^2}}$	Root mean square of ST
IBI	$IBI^{\wedge x}$	Minimum of IBI	ST	ST^{a}	Y intercept of regression line
ST	ST^{b}	slope of regression line			

4.1.3 Processing of Salivary Cortisol

The saliva samples are processed to obtain cortisol concentration values in $\mu g/dL$. The assay range of the cortisol concentration values is between 0.012-3 $\mu g/dL$. Salivary cortisol samples were collected a total of 5 times during the study at time points T1, T2, T3, T4, and T5 (Figure 4.2). To categorize the cortisol concentration values as stressed (S) or not-stressed (NS) class, the following preprocessing steps were implemented.

- Samples at time point T1 and T2 (X_{T1} and X_{T2}) were averaged and subtracted from the remaining three samples (X_{T3}, X_{T4}, X_{T5}) at time points T1, T2 and T3 respectively.

- The three samples, X_{T3}, X_{T4}, X_{T5} from all participants were integrated and standardized.

- A sample X_{Ti} is classified as stressed (S) class if $X_{Ti} > \bar{X}$ and was classified as not-stressed (NS) if $X_{Ti} <= \bar{X}$ (i=3,4,5), where $\bar{X}$ is the population mean.

Table 4.2: Mean and Standard Deviation of cortisol concentration during each timestamps (© 2021 IEEE).

TimeStamps	**Mean (ug/dL)**	**SD (ug/dL)**
T1	0.185	0.138
T2	0.189	0.133
T3	0.172	0.105
T4	0.154	0.078
T5	0.137	0.069

The statistics for stress class are 0.508 ± 0.413 and that of not-stressed class are -1.01 ± 1.05 with p-value $2.15e^{-20}$. The mean and the standard deviation of the cortisol concentration for each of the five time points are shown in Table 4.2

Calibration of Salivary Cortisol with Physiological Signals

Salivary cortisol measurements were obtained during our year-long experiment from about 40 older adults. The cortisol samples were collected during the TSST (Trier Social Stress Test) protocol [93]. The protocol consists of a waiting period, a pre-stress period, stress period, and recovery periods (Figure 4.2). More details on the TSSST protocol is provided in Section 4.2.3. Salivary cortisol collected from saliva samples typically has a lag of 20 minutes from the point of acquisition [102]. Hence, it is important to calibrate the physiological signal with its corresponding cortisol value. The calibration

is performed by shifting the cortisol concentration values 20 minutes to the left to account for the delay in cortisol response.

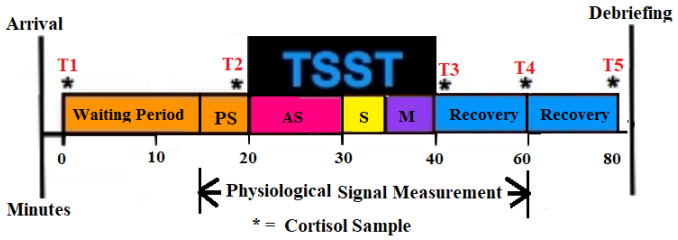


Figure 4.2: Experimental Protocol. Cortisol samples are collected at time points T1, T2, T3, T4, and T5 in 20 minutes duration. Cortisol concentration at time point T_i corresponds to the stress level at time point T_{i-1} (© 2021 IEEE).

For example, in Figure 4.2, cortisol concentration collected at time point T2 corresponds to the stress level at time point T1. Similarly, the cortisol concentration collected at time point T3 is actually the stress level at time point T2 and so on. This calibration technique is used for labeling a particular section as either stressed or not-stressed as evaluated from cortisol concentration values.

4.1.4 Training and Testing

After extracting features from respective feature extraction units, the features are aggregated and annotated with the stress labels obtained from the processing of salivary cortisol. The feature set is first split into train and test set in the approximate ratio of 75-25. The splitting is done such that no two samples in the train and test set should come from the same partic-

ipant. The p-value for the train and test set is $1.73e^{-17}$. Subsequently, a feature selection procedure is applied on the training set to select the best set of performing features from EDA, BVP, IBI, and ST signals. The objective of the feature selection is to evaluate the correlation of the features with stress labels obtained from salivary cortisol concentration and thereby select those features which are statistically significantly correlated with the target variable. The training feature set is correlated with the stress labels obtained from cortisol concentration using Kendall's tau correlation [103]. The correlation coefficient (ρ) and p-value pair are obtained by computing the feature-wise correlation. Only the features with p-values less than 0.05 have been selected. This is because features whose p-value with the target variable is less than 0.05 provide strong evidence against the null hypothesis and hence the alternate hypothesis that the particular feature is strongly correlated with the target variable can be accepted with a high degree of confidence. Figure 4.3 shows the taxonomy diagram of the selected features from each of the signal streams.

We can see from Figure 4.3, that EDA peak amplitude and peak width are found to be important quantities with each having 6 and 4 statistical measure respectively that correlates significantly with the target variable along with minimum prominence of EDA peak. For BVP peak, amplitude and prominence of the peak are found to be the most important quantities with 3 and 6 statistical features respectively. Among the statistical features from peak width, only the median peak width is found to be useful. Furthermore,

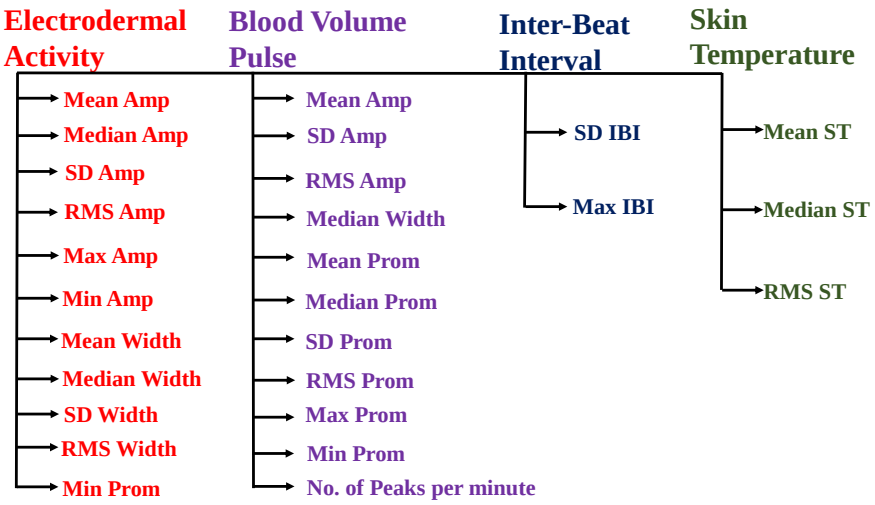


Figure 4.3: Selected features from EDA peak, BVP peak, IBI and ST signal streams (© 2021 IEEE).

the number of peaks occurring per minute was also found to be significantly correlated. From the IBI signal, the standard deviation and maximum of IBI are found to be significantly correlated. Finally, for ST signal, mean, median and root mean square is found to be significantly correlated with the target variable. These final sets of features, 11 features from EDA and BVP signal, 2 from IBI, and 3 from ST constitute the final set of features that will be used for training and testing the machine learning classifier. In the next section, we will present the experimental setup adopted for this study.

4.2 Experimental Setup

In this section, we will discuss the set-up for the experimental study in the context of the inclusion criteria of participants, physiological signal recording,

collection of salivary cortisol, and the protocol for the experiment.

4.2.1 Participant Inclusion Criteria

40 healthy older adults between the age range 60-80 were selected for this study. The participants were screened for existing heart-related conditions, post-traumatic disorder, anxiety disorders, unstable angina, Addison's disease, or Cushing disease before enrolling for the study. Participants were also excluded if they had trouble following instructions, trouble with daily activities, thinking problems, impaired consent capacity, and diagnosis of dementia.

4.2.2 Physiological Data Recording

Physiological signals were collected using a commercially available smart wristband device that has EDA, PPG, and ST sensors embedded in it. Four types of signals were estimated from the three sensors. EDA signal is recorded using two AgCl electrodes placed at the lower wrist of the user. The PPG sensor consists of a green and red light source and a light receiver. PPG sensor outputs the BVP signal using a proprietary algorithm by combining the lights reflected from both green and red light exposure. Another propriety algorithm is used to remove wrong beats from a BVP signal to output the IBI signal. Finally, the temperature signal is collected using a temperature sensor that uses the principle of infrared thermopile to measure the skin

temperature. The sampling rate of EDA, PPG, and ST sensors embedded in the wristband is 4 Hz, 64 Hz, and 4 Hz respectively. These signals can be collected either in recording mode, where the data gets stored in in-device memory and can be later synced to cloud storage for processing, or in the streaming mode where the data is transmitted via a low-energy Bluetooth connection to an edge-interface, such as a mobile application and can be later synced to cloud at the end of a session. In our experiment, we have collected the data in recording mode.

Saliva samples were collected a total of 5 times during the study at time points T1, T2, T3, T4, and T5 (Figure 4.2). Saliva samples have been collected using a cotton swab placed under the participant's tongue for about 2 minutes. The first sample, at time point T1, is collected when the participant first arrives at the study facility. The second sample at time point T2 has been collected after 20 minutes from T1 and at the end of the PS (Pre-Stress) period. The third sample is collected after the stress phase, which is at time point T3. The fourth sample is collected at time point T4 which is after the first recovery phase and the last sample is obtained at time point T5 which is after the second recovery phase.

4.2.3 Experimental Protocol

The TSST (Trier Social Stress Test) protocol has been adopted to induce stress in our study. TSST is a well established experimental protocol known for its ability to induce stress in a naturalistic environment [93]. The phases

involved in the experimental protocol are illustrated in Figure 4.2 (Section 4.1.3).

The TSST protocol consists of a waiting period, pre-stress period, stress period, and recovery period. During the waiting period, the participants are asked to complete demographic surveys and get clarified about any questions about the study process they may have. Following the waiting period is the pre-stress period during which baseline measurements are taken. The waiting period and pre-stress period together is of 20 minutes duration (T1-T2). The pre-stress period is followed by the stress period which is of 20 minutes duration (T2-T3). The stress period is divided into three parts as anticipatory stress (AS) period which is for 10 minutes, followed by subsequent speech (S) and mental math (M) tasks for 5 minutes each. During the AS phase, the participants are shown a topic on which they have to speak continuously for 5 minutes while being observed. During the mental math task, participants are required to solve simple addition and subtraction problems while being observed. The difficulty level of the mental math questions increases with every correct answer and it's mandatory for the participants to answer each question correctly before moving on to the next question. Finally, the study concludes with two recovery phase of 20 minutes each (T3-T5).

4.3 Results and Analysis

In this section, we will report the results of the performance of the machine learning classifier to distinguish between stressed and not-stressed states. Random forest classifier is chosen as the machine learning model to be used to classify between the two states. We will begin this section by first stating the objective of our evaluation and defining the metrics we will be using to evaluate the performance of our classifier.

4.3.1 Evaluation objective and performance metrics

Table 4.3: Performance metric for different sensor combination (© 2021 IEEE).

Sensor Combination	**Signal Combination**	**Total Feature**	**Selected Feature**	**F1-score Stressed**	**F1-score Non-stressed**	**Macro F1-score**	**Accuracy** (%)
EDA	EDA	18	11	0.88	0.79	0.83	84
EDA,PPG	EDA,BVP	35	22	0.91	0.81	0.86	88
EDA,PPG	EDA,BVP,IBI	41	24	0.93	0.85	0.89	90
EDA,PPG,ST	EDA,BVP,IBI,ST	47	27	0.95	0.90	0.92	94

The objective of our evaluation is to quantify the effectiveness of fusing different signal streams coming from multiple sensors in improving the classification rate. To evaluate the performance of the classification rate, the first metric that we will be using is the F1-score for both the individual classes, that is F1-score for stressed class and that of not-stressed class. F1-score is the harmonic mean of precision and recall and is a useful measure of evalu-

ating performance especially when there is an uneven class distribution and also when both false positives and false negatives are equally expensive. The second metric is Macro-average F1-score. This particular metric provides an overall performance estimate for both the classes while giving equal weight to both the majority and minority classes. Finally, we will be using the area under the curve (AUC) score computed using the receiver operator characteristic (ROC) curve to evaluate the performance of the classifier taking different combinations of signal streams. The ROC curve is generated by plotting the true positive rates and false positive rates obtained by varying the threshold used to separate the two classes based on the log probabilities of the score for each prediction. The AUC score is the area under the ROC curve and is a useful measure to quantify the performance of a classifier. All these metrics will be used in conjunction with each other for a detailed analysis of the classification performance.

4.3.2 Performance analysis

The performance of the combination of EDA, EDA-BVP, EDA-BVP-IBI, and EDA-BVP-IBI-ST is shown in Table 4.3. The ROC curve for the same combinations in shown in Figure 4.4. The dotted red line show the ROC curve when random guessing is used to classify between the two classes, The blue, yellow, green and red curves shows the ROC curve of the combination of EDA, EDA-BVP, EDA-BVP-IBI and EDA-BVP-IBI-ST respectively.

While using only the EDA signal, the F1-score for the stressed class is

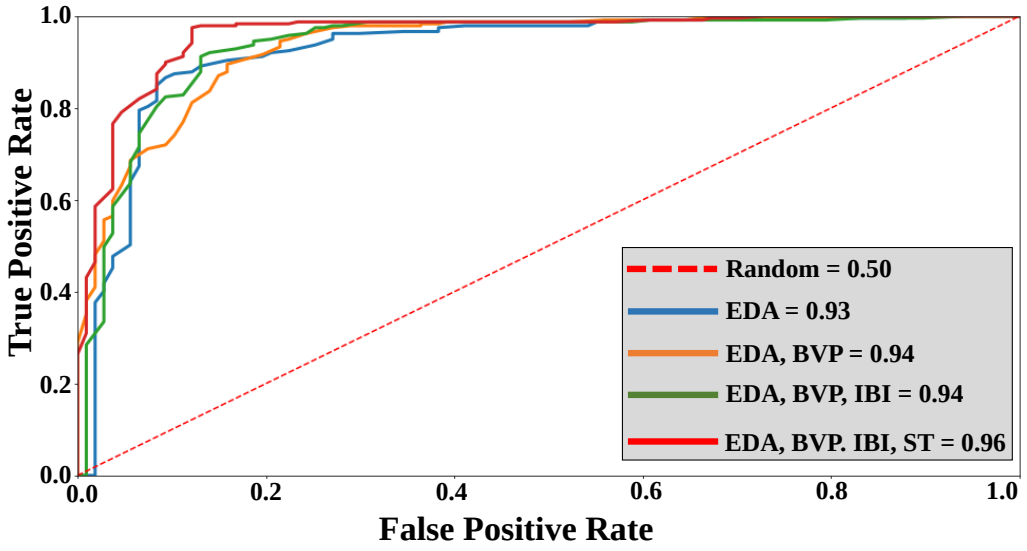


Figure 4.4: Plot of ROC curve for different signal combination (© 2021 IEEE).

found to be 0.88. However, for the not-stressed class, the F1-score is only 0.79 and the macro-average F1-score is 0.83 and the overall accuracy is 84%. Hence, EDA peaks which are indications of emotional arousal can detect the occurrence of stressed class with significantly good accuracy, but also increases the likelihood of mistakenly classifying several non-stressed states as a stressed state.

When the PPG sensor is combined with the EDA sensor and EDA and BVP signal streams are used in conjunction, the F1-score for the stressed class is 0.91, and that of the not-stressed class is 0.81 which is a 3.4% and 2.53% increase respectively when only EDA signal was used. The macro average score and overall accuracy are 0.86 and 88% respectively which is a

3.6% and 4.8% increase than using only EDA signal. Also, it can be seen from Figure 4.4, that the area under the curve for EDA is 0.93 and that for EDA-BVP combination is 0.94.

When the signal streams EDA, BVP, and IBI are combined, a 2.19% increase in the F1-score of the stressed class is observed, and about a 5% increase in the F1-score of the not-stressed class is observed. The macro average F1-score is also increased from 0.86 to 0.89 and the overall accuracy increased from 88% to 90% which are 3.4% and 2.2% increase respectively.

Finally, when all the signal streams are combined, that is when using a combination of EDA, BVP, IBI, and ST signal, we can see an increase of 2.15% from 0.93 to 0.95 in F1-score of stressed class. The F1-score of not-stressed increased from 0.85 to 0.90 which is a 5.88% increase. The Marco-average F1-score increased from 0.89 to 0.92 and the total accuracy increased to 94% which is a 4.44% increase. Further, the AUC score also increased from 0.94 to 0.96.

The improvement of the classification performance by combining the four different signal streams can be visualized in Figure 4.5. The stressed class is represented as 1 and the not-stressed class as 0 and is plotted as a step function with respect to time for the participants in the test set. The actual classification of the stress levels obtained from cortisol concentration is shown in red. The predictions using EDA, EDA-BVP, EDA-BVP-IBI, and EDA-BVP-IBI-ST are shown in blue, green, magenta, and black respectively. The number of misclassification when using only the EDA signal is 54. The

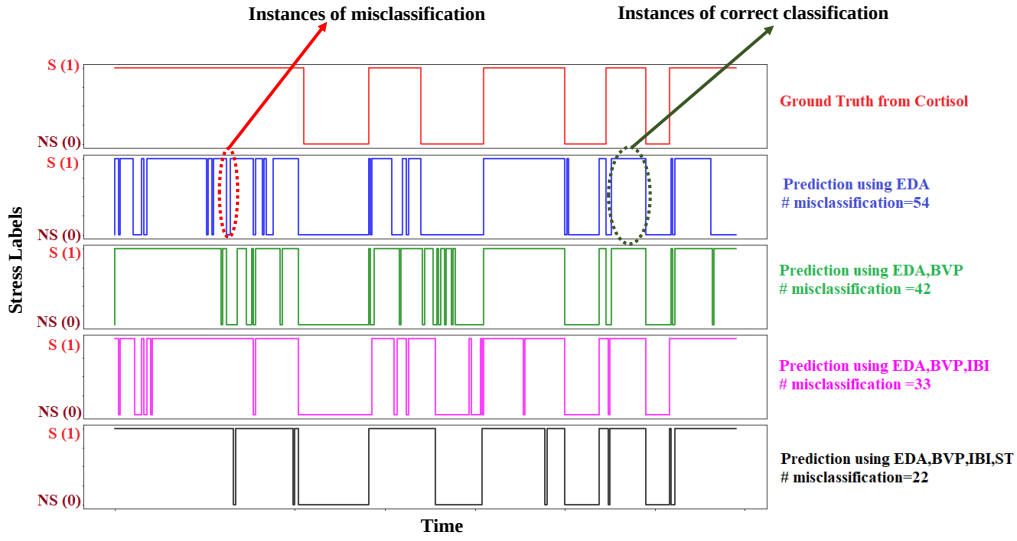


Figure 4.5: Prediction using the four signal combination with ground truth referenced from salivary cortisol. Stress class represented as 1 and not-stressed class as 0 (© 2021 IEEE).

number of misclassifications reduced to 42 when a combination of EDA and BVP is used. Using a combination of EDA, BVP, and IBI further reduced the misclassification to 33, and to 22 when all the four signal streams EDA, BVP, IBI, and ST are combined.

4.4 Voice-Query Based Prototype Framework for the Proposed Stress Detection System

In this section, we will discuss the voice-query based framework, prototyped for the integration of the proposed stress detection model in consumer end devices. The user can query for their vitals and stress levels using a voice interface such as a mobile application, smart speaker, or any smart consumer devices with voice assistant capabilities. The realization of the voice-based framework is visually represented in Figure 4.6.

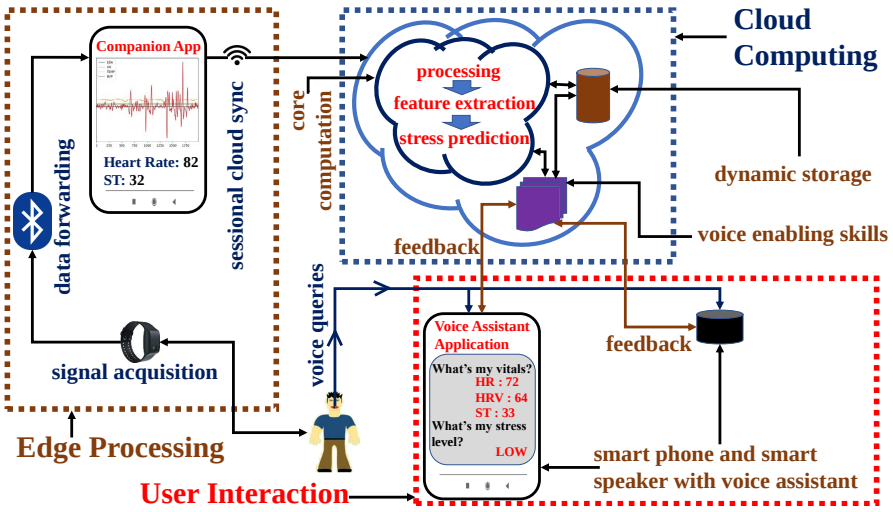


Figure 4.6: Prototype framework for the integration of the proposed stress detection system in a voice based consumer end device such as smart speaker or mobile application. The user can initiate a voice query on either of these devices for feedback on their vitals and stress levels (© 2021 IEEE).

The prototype consists of three main modules: (i) Edge Processing, (ii)

Cloud Computing, and (iii) User Interaction. In edge processing, the smart wristband records EDA, BVP, IBI, and ST signals using EDA, PPG, and ST sensors, and transmits these signals using a low energy Bluetooth connection to a companion app on a mobile phone. The companion app performs simple computations on the individual signal streams such as visualizing the signal plots and calculating instantaneous heart rate from the IBI signal. The signal streams are then transferred to an existing cloud resource after every session through a WiFi connection for stress level prediction. The cloud computing module performs three main functions: (i) Core computations, (ii) Storage, and (iii) Hosting voice enabling skills. The core computation consists of the processing of signal streams, feature extraction, and stress level prediction. The storage module is used for storing the intermediate data for faster processing and as well as could store historical data for future analysis. The third component is the voice enabling skills that communicate with the cloud processing module, the storage module, and user endpoints such as a smart speaker or mobile application for receiving the query, processing them, and providing the requested feedback.

The user interaction module consists of a mobile phone that hosts the companion app and the voice assistant application through which the user can obtain feedback on the vitals. A smart speaker could also be used as an endpoint for querying and receiving voice-based feedback. In our prototype framework, the integration of a voice interface as a means to query and get feedback is validated by simulating voice dialogues on a voice-based web

application hosted on a mobile phone. The voice-based web application is capable of handling requests for checking both the vitals like heart rate, heart rate variability, and skin temperature, and stress level. We have analyzed the feasibility of the proposed prototype in transmitting multiple signals from a wristband in real-time to the companion app and subsequently to the cloud resource. This was done by simulating a data streaming scenario, where a user wore the smart wristband and the signals were streamed continuously for one and a half hour. After the session ended, the battery of the smart wristband is only reduced by a small amount, and the companion app consumed a negligible percent of the phone's battery. Moreover, the wristband used in the experiment can stream data continuously for about 20 hours in streaming mode, and record data continuously for about 30 hours in recording mode. This shows the feasibility of the proposed system for realizing a low-power consumer electronic system for monitoring stress using only a smart wristband.

4.5 Conclusion

In this work, four physiological signals, EDA, BVP, IBI, and ST signals from three wrist-worn sensors, EDA, PPG, and ST were used to develop a stress detection model for older adults. The four physiological signals along with salivary cortisol measurement were recorded from 40 older adults (28 females and 12 males) and age 73.625 $\pm$ 5.39 during the TSST protocol. Features

characterizing the amplitude, width, and prominence of EDA and BVP peaks were extracted along with statistical features from IBI and ST signal. A total of 47 features were extracted from the four signals, out of which 18 were from EDA, 17 from BVP, 6 from IBI, and 6 from ST. Out of the 47 features, 27 features were found to be significantly correlated with the target variable. Out of the 27 features, 11 were from EDA and BVP each, 2 from IBI, and 3 from ST. A combination of these features was used to train and test a random forest classifier to distinguish between the two states. Results show that the overall accuracy increased by 4.8% when a combination of EDA and BVP was used as opposed to using only EDA. The accuracy further increased by 2.2% when features from IBI were combined which further increased by another 4.4% when features from ST were also combined along with EDA, BVP, and IBI. The fact that the prediction accuracy of the classifier increases when sensor fusion is performed shows the promise of the proposed method over simpler techniques such as estimating stress using just heart rate or only one physiological parameter. Further, a voice-based consumer use case is prototyped, where the user can query for vitals and stress levels using a voice interface like a mobile phone or smart speaker.

The results and analysis presented in this work show its merit in being a suitable choice for designing a stress detection model for older adults. Leveraging the use of such wrist-worn sensors to detect stress from physiological signals by correlating with cortisol, which is a clinically accepted biomarker for stress can improve the state-of-the-art stress detection model. This is

because using cortisol as the reference to detect stress can bring clinical level diagnosis of stress into the everyday consumer world and hence can improve the quality of consumer health care. Such a model can be used for continuous stress monitoring in an unobtrusive way, thereby enabling individuals to manage stress by themselves which would promote well being and improve the quality of life for older adults. As future work, the proposed method for stress detection can be extended to design a stress monitoring and management system using the prototype voice-based framework. The prototype framework can be used to monitor stress levels unobtrusively and on a continuous basis, and additional functionality could be added to the existing prototype for launching personalized intervention for stress reduction.

Chapter 5

Anxiety Detection Using Wearable Sensor and Context Feature

In the previous two chapters, we presented our work on detecting stress using fingertip-based wearable sensors and wristband-based wearable sensors. In Chapter 4, we have seen how fusing features from multiple signal streams increases the performance of the predictive model for detecting stress. Similarly, as discussed in Chapter 2, Section 2.5.2, it has been observed that fusing features from multiple signal streams increases the performance in detecting anxiety. Although fusing different physiological features increases the performance of the anxiety detection model, the complexity of integrating and processing multiple physiological signals in real-time could be a significant

design challenge.

In this chapter we will present the work on detecting anxiety using a single wearable physiological sensor and context feature. The proposed method for anxiety detection combines features from a single physiological signal with an experimental context-based feature to improve the performance of the anxiety detection model. The experimental data for this work has been obtained from a year-long experiment on 41 healthy older adults (26 females and 15 males) in the age range 60-80 with mean age 73.36 $\pm$ 5.25 during a TSST (Trier Social Stress Test) protocol. The anxiety level ground truth was obtained from STAI (State-Trait Anxiety Inventory), which is regarded as the gold standard to measure perceived anxiety. EDA and BVP signals were recorded using a wrist-worn EDA and PPG sensor respectively. The phases of the experimental study are encoded as unique integers to form the single column context feature vector. A combination of features from a single sensor with the context feature vector is used for training a machine learning model to distinguish between anxious and not-anxious states. Further, end-to-end processing of EDA and BVP signals was simulated for real-time anxiety level detection.

This chapter is organized as follows, Section 5.1 will present the proposed work for anxiety detection. Section 5.2 will present the experimental protocol adopted for the study. Section 5.3 will present the discussion on the results and analysis, and finally Section 5.4 will conclude the chapter.

5.1 Proposed Method for Anxiety Detection

This section will describe the proposed method for anxiety detection. The proposed method for anxiety detection fuses an experimental context feature with features from a single physiological signal to train a machine learning classifier for accurately distinguishing between anxious and non-anxious states. The overview of the proposed method for anxiety detection is shown in Figure 5.1.

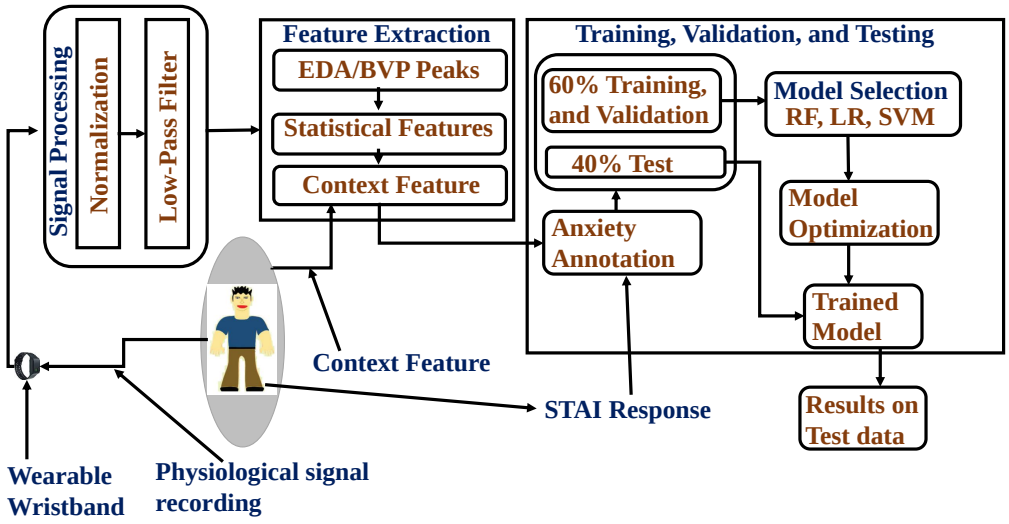


Figure 5.1: Overview of the proposed method for anxiety detection. RF=Random Forest Classifier, LR=Logistic Regressiom, and SVM=Support Vector Machine. STAI=The State-Trait Anxiety Inventory

The major computational blocks of the proposed methods are: i) Signal Processing, (ii) Feature Extraction, and (iii) ML (Machine Learning) Modeling.

5.1.1 Signal Processing

This computational unit preprocesses the raw signal before the feature extraction stage. In this unit, the EDA and the BVP signals are first normalized by scaling the sensor values in the range [0,1]. After normalization, a low-pass Butterworth filter of order 5 was used to remove high-frequency components from the signal, which usually occurs as a result of motion artifact. For the EDA signal, the signal components that were higher than 1 Hz were cut off and for the BVP signal, the signal components that were higher than 10 Hz were cut off.

5.1.2 Feature Extraction

In this section, we will describe the physiological features extracted from EDA, and BVP signals and the context feature used for training the machine learning model. For analysis, 3 minutes of data from pre-stress period, and about 15 minutes of data each from stress and relax phase were used. The physiological features are extracted from physiological signals using a running window of 30 seconds and an overlap of 15 seconds. On an average, 118 data points were generated from each participant and for 41 participants, a total of 4853 data points were obtained. The notation of the extracted physiological features from the three physiological signals are tabulated in Table 5.1

Feature Extraction from EDA Signal

EDA signal measures the skin conductance of the skin which varies according to the sweating produced by the sweat glands. Arousal due to stressful situations results in the increased activity of the SNS (Sympathetic Nervous System) which leads to various physiological changes [104]. Sweating is one of the physiological changes that is modulated by the SNS. Hence we hypothesize that using features from the EDA signal can be useful in detecting the onset of anxiety caused due to stressful events. Previous research has shown that the peak characteristic of the EDA signal is useful in detecting stress levels of an individual [73] [51]. Hence, we hypothesize that perceived anxiety due to emotional arousal resulting from stressful situations could also be modeled using the peak features from EDA. Peak detection algorithm [97] is used to compute the amplitude, width, prominence of peaks occurring in a given window. Subsequently, statistical measures such as the mean, standard deviation, median, root mean square, minimum, and the maximum of peak amplitude, width, and prominence are computed to form the final set of 18 features.

Feature Extraction from BVP Signal

The BVP signal is a measure of cardiovascular activity. BVP signal can be an indicator of cardiovascular arousal during exposure to stressful situations [98]. Previous research that has used cardiovascular arousal as a measure to detect the onset of anxiety has resulted in satisfactory results. However, most

of the existing research has used ECG to measure cardiovascular arousal in the context of anxiety detection. Since BVP-based features were successfully used previously in the context of stress detection [51], we hypothesize that BVP-based features can also be used to detect the onset of anxiety occurring as a result of stressful situations. Similar to EDA, peak detection algorithm [97] is used to compute the width, amplitude, and prominence of the systolic peaks occurring in a given window. Along with statistical measures, the frequency of systolic peaks per minute is also computed resulting in a total of 17 features.

Table 5.1: Notation and Description of the extracted features from EDA, and BVP.

Signal Source	Feature Notation	Feature Description	Signal Source	Feature Notation	Feature Description
EDA	$P_{amp}^{\bar{x}}$	Mean of peak amplitudes	BVP	$S_{min}^{\#x}$	No. of peaks per minute
EDA	$P_{amp}^{\tilde{x}}$	Median of peak amplitudes	BVP	$S_{width}^{\bar{x}}$	Mean of peak widths
EDA	P_{amp}^{σ}	Standard deviation of peak amplitudes	BVP	$S_{width}^{\tilde{x}}$	Median of peak widths
EDA	$P_{amp}^{\sqrt{\bar{x^2}}}$	Root mean square of peak amplitudes	BVP	$S_{width}^{\sigma x}$	Standard deviation of peak widths
EDA	$P_{amp}^{\vee x}$	Maximum of peak amplitudes	BVP	$S_{width}^{\sqrt{\bar{x^2}}}$	Root mean square of peak widths
EDA	$P_{amp}^{\wedge x}$	Minimum of peak amplitudes	BVP	$S_{width}^{\vee x}$	Maximum of peak widths
EDA	$P_{width}^{\bar{x}}$	Mean of peak widths	BVP	$S_{width}^{\wedge x}$	Minimum of peak widths
EDA	$P_{width}^{\tilde{x}}$	Median of peak widths	BVP	$S_{prom}^{\bar{x}}$	Mean of peak prominence
EDA	P_{width}^{σ}	Standard deviation of peak widths	BVP	$S_{prom}^{\tilde{x}}$	Median of peak prominence
EDA	$P_{width}^{\sqrt{\bar{x^2}}}$	Root mean square of peak widths	BVP	$S_{prom}^{\sigma x}$	Standard deviation of peak prominence
EDA	$P_{width}^{\vee x}$	Maximum of peak widths	BVP	$S_{prom}^{\sqrt{\bar{x^2}}}$	Root mean square of peak prominence
EDA	$P_{width}^{\wedge x}$	Minimum of peak widths	BVP	$S_{prom}^{\vee x}$	Maximum of peak prominence
EDA	$P_{prom}^{\bar{x}}$	Mean of peak prominence	BVP	$S_{prom}^{\wedge x}$	Minimum of peak prominence
EDA	$P_{prom}^{\tilde{x}}$	Median of peak prominence	BVP	$S_{amp}^{\bar{x}}$	Mean of peak amplitudes
EDA	P_{prom}^{σ}	Standard deviation of peak prominence	BVP	$S_{amp}^{\sigma x}$	Standard deviation of peak amplitudes
EDA	$P_{prom}^{\sqrt{\bar{x^2}}}$	Root mean square of peak prominence	BVP	$S_{amp}^{\sqrt{\bar{x^2}}}$	Root mean square of peak amplitudes
EDA	$P_{prom}^{\vee x}$	Maximum of peak prominence	BVP	$S_{amp}^{\vee x \Delta \wedge x}$	Range of peak amplitudes
EDA	$P_{prom}^{\wedge x}$	Minimum of peak prominence			

Context-Based Feature

The importance of context-awareness in the context of anxiety detection was studied in [105]. The elicitation of anxious reaction due to stressful events is a subjective perception and is significantly affected by the context

of the event [106]. Hence, alongside features from physiological signals, using context-based features to train a machine learning model can result in better performance. For example, EDA measures emotional arousal and BVP measures cardiovascular arousal. However, arousals can be positive (because of positive stress or eustress), and negative (because of negative stress or distress). Hence, in situations where an individual is under positive arousal will not perceive the situation as an anxious situation. However, a machine learning model that models anxiety simply based on arousal might result in several false positive.

We hypothesize that encoding context information as a feature variable with the physiological features can help to distinguish the positive arousal from the negative arousal. To implement this, we encoded an experimental context feature with respective experimental segments of the physiological features. There are three distinct experimental context used in our study. The first experimental context is during the PS (Pre-Stress) period, during which the user is not subject to any stimulus (Figure 5.2). The second experimental context is during the AS (Anticipatory Stress), S (Stress), and M (Math), during which the user is subjected to stressful stimulus. The third context is during the Recovery period, during which the user is subjected to relaxation-based stimulus. The context feature used in our system is encoded as -1 for the experimental session when the user was not exposed to any kind of stimulus. The experimental session when the user is exposed to stressful stimulus is encoded as 0. Finally, the session when the user is exposed to a

relaxation-oriented stimulus is labeled as 1. More details on the experimental protocol is described in Section 5.2.3.

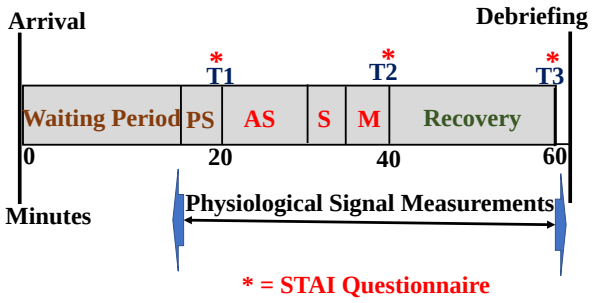


Figure 5.2: Experimental Protocol

5.1.3 Training, Validation, and Testing

This section will describe the machine-learning-based modeling of anxiety. Primarily we will detail the training and testing protocol of the anxiety detection model adopted in this study. The three main components required for training and testing a machine learning model to distinguish between anxious and non-anxious states are the features from physiological signals, the experimental context feature, and the anxiety level ground truth. After the feature extraction, the feature samples are annotated with their respective anxiety states such as 'Anxious' or 'Not-Anxious' states. These anxiety states are determined from the STAI questionnaire filled by the user after each phase of the experimental protocol. Details on how anxiety states are determined from the STAI questionnaire will be described in detail in Section 5.2.2.

For modeling the anxiety states with physiological features and the context feature, we have adopted a supervised machine learning approach. For model evaluation, the entire feature set is divided into a training set (50%), a validation set (10%), and a test set (40%). The training and the validation set will be used to select the appropriate machine learning algorithm for anxiety states classification and hyperparameter tuning. The best performing machine learning algorithm in the training and validation phase will be chosen for testing on the test set.

5.2 Experimental Setup

The objective of our experiment is to model the physiological changes resulting because of stressful situations with perceived anxiety levels for older adults. In this section, we will discuss the inclusion criteria for participants in our experiment, the data recording, ground truth estimation, and finally the experimental protocol adopted for the study.

5.2.1 Participants

The participants recruited for this experiment were 41 older adults (26 females and 15 males) in the age range 60-80 with mean age 73.36$\pm$5.25. Before recruiting for the experiment, the participants were screened for certain existing conditions which might interfere with the results of the experiment. These are heart-related conditions, post-traumatic disorder, anxiety disor-

ders, unstable angina, Addison's disease, or Cushing disease. Other than the existing conditions, participants were not recruited if they had any trouble with daily activities, trouble following instructions, impaired consent capacity, thinking problems, and have a diagnosis of dementia.

5.2.2 Data Recording

During the experiment, three different types of data were recorded: (i)physiological signals from wristband sensors, (ii) response to STAI questionnaire from the user, (iii) context-based feature from the experimental protocol.

Recording of Physiological Signals

In our proposed anxiety detection, we have recorded two physiological signals during the experimental protocol. EDA signal is recorded with the help of two AgCl electrodes which maintains contact with the lower wrist of the user. The AgCl electrodes are attached to the strap of the wristband. The BVP signal is measured using the PPG sensor which uses a combination of red and green light for estimating the blood volume. The EDA and BVP signals were collected at a sampling rate of 4 Hz and 64 Hz respectively.

The wristband device used in our study has a form factor of 110-190 mm and weighs around 25g. The wristband device executes the data transfer using low energy bluetooth (streaming mode) and has a flash memory (for recording mode) that can record upto 60 hours of data. In our experiment, the recording of EDA and BVP signals were conducted in recording mode

to ensure collection of better signal quality for analysis. These signals were stored in in-device memory and were later synced to a desktop machine through the proprietary cloud interface after the end of each experiment. The collected physiological data did not contain any personal information about the participants.

Recording of STAI Response

The STAI (The State-Trait Anxiety Inventory) is a commonly used self-feedback form to measure state and trait anxiety [107]. The STAI is regarded as the gold standard for measuring anxiety [108][109][110]. The participants were instructed to digitally respond to the 20 items STAI questionnaire intended to quantify state anxiety. Each of the participants was required to fill up the STAI questionnaire during the time points T1, T2, and T3 (Figure 5.2). The STAI questionnaire contains question fields such as "I feel calm", "I feel upset" etc and each of these fields is weighted on a scale of 1-4. The participants were asked to weigh each field according to how much they agree/disagree with the question fields. For example, a high level of agreement with a particular question field, "I feel calm" will be weighted with 4 and a high level of disagreement for the same field will be weighted as 1.

Scoring is done by summing all the scores of all fields. A higher score indicates a higher level of anxiety. Before scoring, the weights of all the question fields that are positive in nature, such as "I feel calm," "I feel

relaxed" are reversed before they are summed. Hence for those fields, a weight of 4 will become 1, 3 becomes 2, and so on. The maximum score possible is 80 and the minimum score possible is 20. The mean and the standard deviation for all participants during the timestamps T1, T2, and T3 are shown in Table 5.2.

Table 5.2: Mean and Standard Deviation of the STAI response during each timestamps for the 41 participants

TimeStamps	**Mean**	**SD**
T1	26.92	8.52
T2	31.54	9.44
T3	25.40	8.52

To classify the score of a particular timestamp as anxious or non-anxious states, scores of all the participants are integrated and standardized. Subsequently, samples whose score is less than the population mean of the standardized score are classified as non-anxious states. Samples whose score is greater than the population mean are classified as anxious states. The statistics of the anxious state ("A") is 1.08 ± 0.801, and that of non-anxious state ("NA") is -0.668 ± 0.2144. The p-value between the scores of anxious and non-anxious states is $5.29e^{-20}$ which indicates a significant statistical difference between the anxious and non-anxious states.

Recording of Context Feature

Since in this work we are evaluating a single contextual feature, the integration of context feature, based on the different phases of the experimental

protocol (discussed in Section 5.1.2) with physiological features is performed manually. The context feature used in this work is predetermined based on the experimental protocol and not on user feedback. In general, we recommend environmental context-based features that can be collected unobtrusively without requiring intervention from the user.

5.2.3 Experimental Protocol

To induce stress-related anxiety, the TSST (Trier Social Stress Test) is adopted because of its ability to induce stress naturally [93]. The entire protocol begins with a waiting period, during which recording of demographic information, signing of the consent form, and briefing of experiment-related tasks. The phases of the experimental protocol after the waiting period can be divided into the pre-stress period, stress period, and recovery period. During the pre-stress period, baseline measurements are taken. The stress period begins with an AS (Anticipatory Stress) period where the participant is required to prepare a 5-minute speech task based on a given topic. The AS period lasts for 10 minutes, after which there is a subsequent speech and math task for 5 minutes each. The purpose of the stress stimulus is to increase the perceived feeling of anxiety in older adults in a natural way. The high mean value of the state scores after time point T2 (Table 5.2) shows that the experimental protocol was successful in inducing anxiety as a result of exposure to stressful situations. The final phase is the recovery phase in which the participants were exposed to a relaxing stimulus intended for the

participants to return to the non-anxious state.

5.3 Results and Analysis

This section will discuss and analyze the results of the proposed anxiety detection system. First, we will analyze the training and validation of the machine learning models in distinguishing between anxious and non-anxious states. We will then perform a qualitative analysis on the best-performing machine learning model in the training and validation phase. Subsequently, we will perform an extensive analysis of the selected model on its performance on the test data. Finally, we will analyze the feasibility of the trained model in classifying between anxious and non-anxious states in real-time.

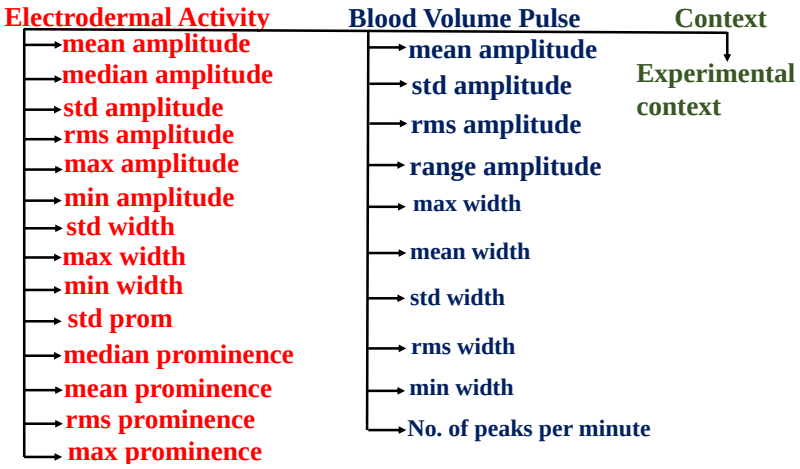


Figure 5.3: Taxonomical representation of the selected features

5.3.1 Training and Validation

The training set, which is 50% of the entire set is first used to select significantly correlated features with the anxiety labels obtained from the STAI score. This is done by taking Kendall's tau correlation [103] between each feature and the anxiety labels. The features which are only significantly correlated (p-value$<$0.05) are selected for training, validation, and testing the machine learning classifiers. Figure 5.3 shows the taxonomical representation of the selected features from each signal stream.

From Figure 5.3, we can see that 14 features out of the 18 features extracted from the EDA signal are found to be significantly correlated with the anxiety labels. The 14 features include the mean, median, RMS, maximum, and minimum of the EDA peak amplitudes along with standard deviation and maximum of EDA peak width and prominence. Other features selected from EDA are the minimum of EDA peak width, median, RMS, and mean of EDA prominence. From BVP signals, only the statistical measures of systolic peak width and amplitude were found to be significantly correlated with anxiety labels together with the frequency of systolic peaks per minute. The experimental context is also found to be significantly correlated with the anxiety labels. Hence, a total of 14 features were selected from EDA, and 10 features from BVP were selected for training the machine learning algorithms.

We have evaluated three machine learning algorithms in the training and validation phase. These machine learning algorithms are random forest clas-

sifier (RF), logistic regression (LR), and support vector machine (SVM). Table 5.3 shows the validation score of the three different machine learning algorithms on the validation set in distinguishing between anxious and non-anxious states. The hyperparameters of the machine learning models are tuned using an exhaustive grid search 5 fold cross-validation.

Table 5.3: Validation scores of different sensor combination with context-feature for different machine learning algorithms. RF=Random Forest, LR=Logistic Regression, SVM=Support Vector Machine, C=Context

Algorithm	**EDA**	**EDA+C**	**BVP**	**BVP+C**
RF	0.88	0.91	0.78	0.82
LR	0.51	0.61	0.53	0.61
SVM	0.37	0.58	0.45	0.61

From Table 5.3, we can see that the RF classifier performed the best among LR and SVM in this case. RF algorithms achieved a validation score of 0.88 with EDA features. When the context feature is fused along with the EDA feature, the validation score increased to 0.91. Similar performance improvement is observed with the BVP signal when context features are fused with those features. Figure 5.4 shows the validation scores obtained by different hyperparameter combinations. From Figure 5.4, we can visualize that for almost all hyperparameter combinations, the combination of context feature and physiological features performs better than using only the physiological features.

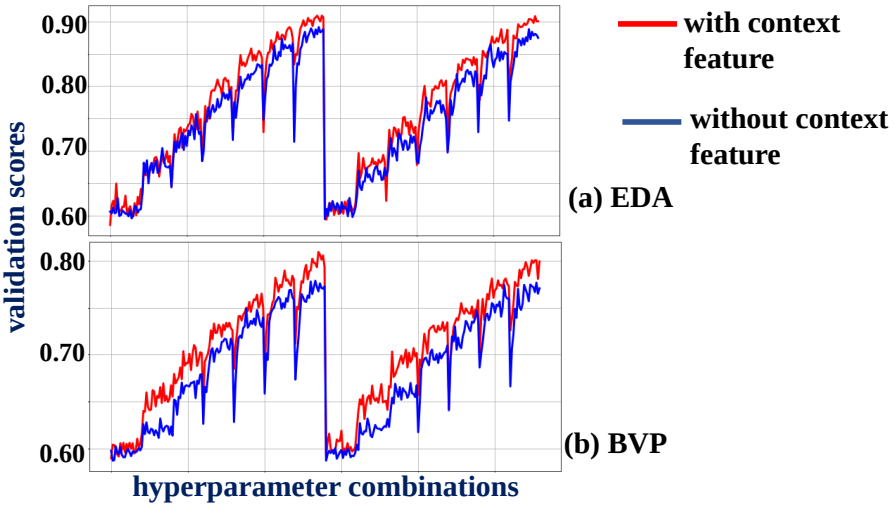


Figure 5.4: Validation scores obtained by different hyperparameter combination for EDA, and BVP signals. The red lines indicate the scores obtained with combination of context feature and physiological features and the blue lines indicate the score obtained with only physiological features.

Although LR and SVM did not perform as well in this dataset, the result is still encouraging because even for those classifiers, improvement in performance is observed when context feature is fused with the features of EDA and BVP signals. For example, LR achieved a 19% increase in performance when the context feature is fused with EDA feature. Similar performance improvement is observed for the BVP signal as well. For SVM, the validation score increased by 56% when the context feature is fused with the EDA feature. This highlights the importance of the context feature for distinguishing between anxious and non-anxious states. Since the RF model performed the best, we will be importing the optimized model for qualitative model evaluation and subsequently for testing and deployment.

5.3.2 Evaluation of the Trained Random Forest Classifier

In this section, we will perform a qualitative evaluation of the trained random forest model. Table 5.4 shows the hyperparameters of the 4 trained models using the 4 feature combinations. Table 5.4 shows the optimized model hyperparameters using cross-validation.

Table 5.4: Optimized hyperparameters for the six trained models using different feature combinations. C represents the context feature

Trained Models	**Split Criterion**	**Maximum Depth**	**Number Estimators**
EDA	Gini	7	20
EDA+C	Gini	7	13
BVP	Gini	7	18
BVP+C	Entropy	7	12

Table 5.5: Performance metrics of the trained models on the test data

Trained Models	**Number of Features**	**F1-score Anxious**	**F1-score Not-Anxious**	**Macro F1-Score**	**Accuracy (%)**
EDA	14	0.86	0.91	0.88	89
EDA+C	**15**	**0.90**	**0.93**	**0.92**	**92**
BVP	10	0.71	0.83	0.77	78
BVP+C	**11**	**0.77**	**0.86**	**0.82**	**83**

The RF algorithm was optimized for the split criterion, maximum depth, and the number of estimators. The split criterion determines the quality of a split. Random forest partitions data into different nodes for decision-making by splitting based on feature values. The quality of the split is determined by either of the two criteria, Gini and entropy. Maximum depth is the maximum depth till which a decision tree is allowed to expand. The number of estimators refers to the number of decision trees in the forest.

From Table 5.4, we can observe that the optimized trained model that fuses context feature with physiological features have significantly less number of estimators than the ones that use only physiological features. For example, the trained model that uses EDA and the context feature has 35% fewer estimators than the trained model that uses only EDA features. A similar decrease in the number of estimators is observed for the trained models that used BVP signal. The trained model that used the BVP and the context feature has only 12 estimators as compared to 18 estimators of the trained model that uses BVP features alone. This highlights the importance of the context feature in reducing the model complexity along with increasing the performance of the model.

5.3.3 Results on Test Data

In this section, we will discuss the performance of the trained models (Table 5.4). Table 5.5 shows the results on the test data. The performance of the trained models are evaluated based on the models' capacity to distinguish between anxious ("A") and not-anxious ("NA") states. The performance metrics are the F1-score of the models' ability to detect the anxious states and not-anxious states, average macro F1-score of overall predictions, and overall accuracy. F1-score is the harmonic mean of precision and recall. Macro F1-score is an estimate of the overall performance estimate of the prediction model for both the classes. Macro average F1-score gives equal weight to both classes and hence this metric is not affected by the presence

of class imbalance.

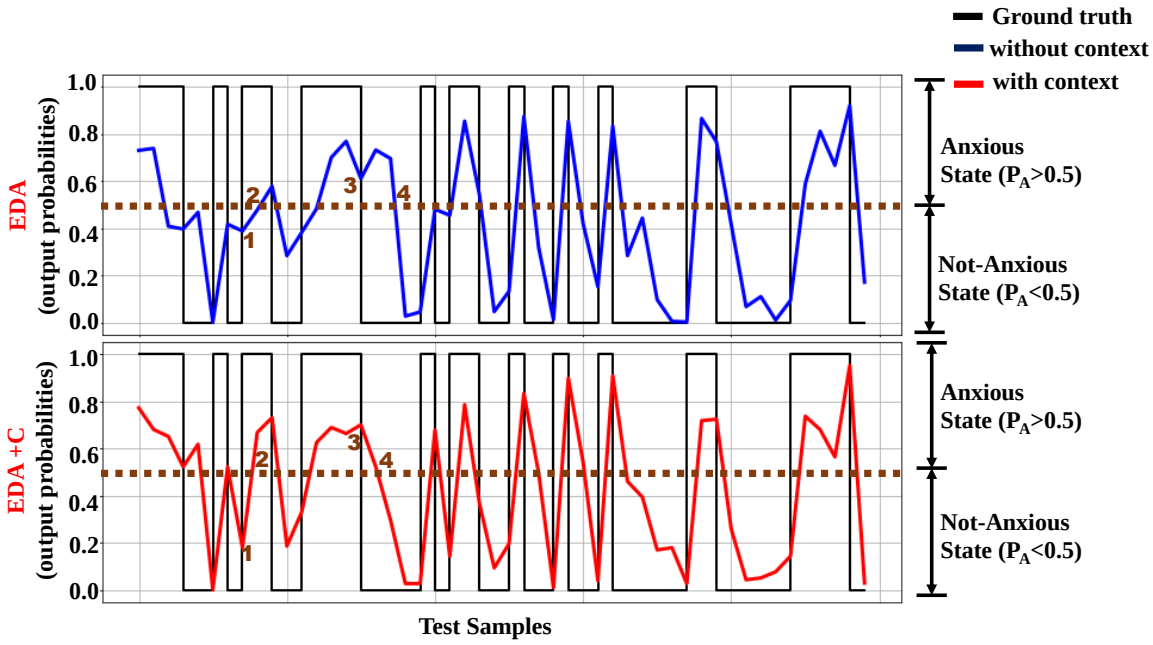


Figure 5.5: Plot of the output probabilities of the machine learning models for EDA and EDA+C models

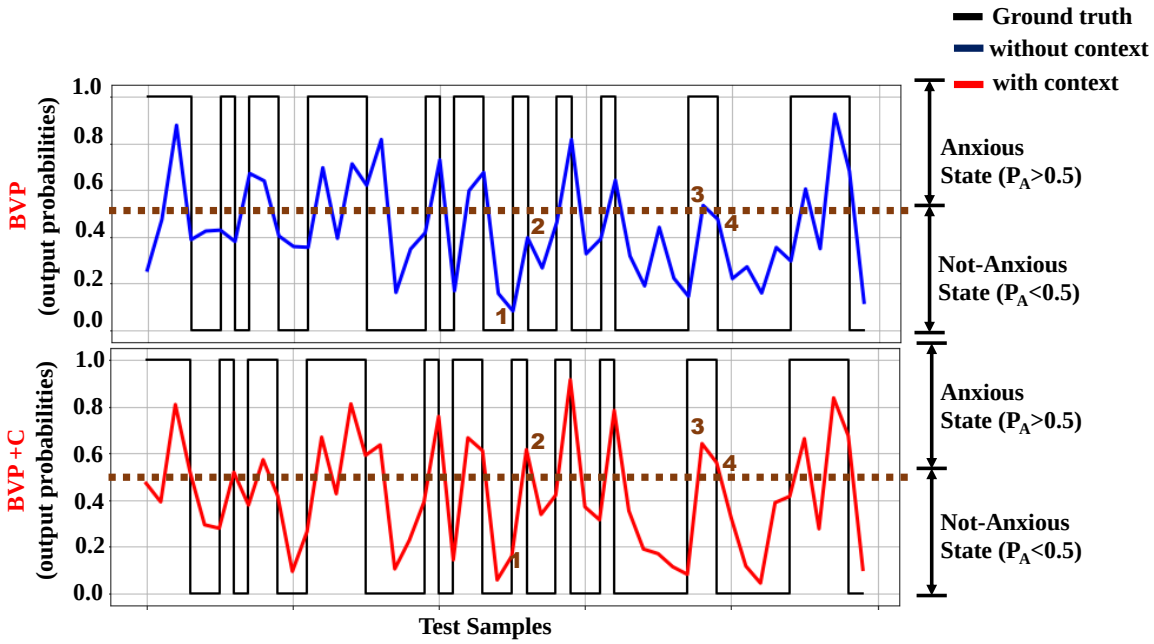


Figure 5.6: Plot of the output probabilities of the machine learning models for BVP and BVP+C models

From Table 5.5, we can observe that both the hybrid physiological-context-based machine learning models-(EDA+C and BVP+C) performed better than the machine learning models that used only physiological signals (EDA and BVP). EDA+C model was able to increase the F1-score of the model in detecting the anxious state and non-anxious states by 4.65% and 2.19% respectively. The macro F1-score and overall accuracy of the EDA+C model were 4.54% and 3.37% higher than that of the EDA model. The BVP+C model achieved an 8.45% and 3.61% higher F1-score for the anxious and non-anxious states respectively than the BVP model. The macro F1-score and overall accuracy for the BVP+C model were also 6.49% and 6.41% higher

than the BVP model respectively.

To further analyze how context-feature can improve the prediction of the machine learning models, we visualize the output probabilities of the machine learning models on the test set. The output probabilities for EDA and EDA+C models are shown in Figure 5.5 and that for BVP and BVP+C are shown in Figure 5.6. To classify a particular feature sample as either anxious or not-anxious, the random forest classifier machine learning models outputs two probabilities, P_A and P_{NA}, where P_A represents the probability of that feature sample to be classified as the anxious state (A), and P_{NA} represents the probability of that feature sample to be classified as the not-anxious state. A threshold (usually 0.5) is set which acts as the decision boundary. Hence, for a particular sample, if $P_A > 0.5$ then the model classifies that sample as anxious. In the same way, If $P_A < 0.5$ implies $P_{NA} > 0.5$, and hence, in this case, the feature sample will be classified as the not-anxious state. In other words, if the probability of the positive class (anxious class in this case) is closer to 1, the sample is classified as anxious and if the probability is closer to 0, the sample is classified as not-anxious.

In Figure 5.5, the output probabilities of some of the test samples are plotted for EDA (top of the figure) and EDA+C (bottom of the figure). The decision boundary (P=0.5) is shown in the brown dotted line. The output of the trained models with context (plotted in red line), and without context (plotted in blue line) is shown along with the ground truth, the actual label (anxious or not-anxious). The ground truth is plotted in a black line as a

Table 5.6: Simulation results on the size of the trained models and latency of the end-to-end processing for real-time anxiety detection

Trained Models	Model Size(KB)	Latency (seconds)				
		Data Loading	Signal Processing	Feature Extraction	Prediction	Total Time
EDA	136	0.004	0.003	0.002	0.009	0.014
EDA+C	109	0.005	0.002	0.003	0.011	0.016
BVP	196	0.037	0.002	0.002	0.013	0.05
BVP+C	115	0.037	0.003	0.002	0.007	0.05

step function. Similar plot is also obtained for BVP and BVP+C models in Figure 5.6.

From Figure 5.5, we can observe that, on addition of context feature, the output probability P_A either decreases or increases to increase the confidence and hence avoid misclassification. For example, in Figure 5.5, for EDA model, we see that the test samples marked between 1 and 2 are mostly classified as ("NA") ($P_A < 0.5$) even though the actual label is ("A"). When the context feature is used along with the EDA features, the output probabilities of some of the test samples are increased above the threshold (0.5) and hence those samples are then correctly classified as ("A"). Similarly, for the test samples in between the region 3 and 4, the actual label is "NA". However the BVP model incorrectly classifies them as "A" ($P_A > 0.5$). When BVP+C model was used, the output probabilities of some of the test samples between the region 3 and 4 reduced below the threshold (0.5) and hence, those samples were correctly classified as "NA".

The same effect of the context-based feature can be visualized in Figure 5.6. For example, all the test samples between the region 1 and 2 are classified

as "NA" ($P_A < 0.5$)by the BVP model. However, when BVP+C model was used, the output probabilities of some of the test samples between the region and 1 and 2, are increased above the threshold (0.5) and hence are correctly classified as "A". Similarly, for test samples between 3 and 4, the output probability of the BVP model is less than or equal to 0.5, even though the actual label is "A". In this case, the classifier may not have enough confidence to classify samples as "A" or might misclassify them as "NA". When BVP+C model is used, the output probabilities increased above the threshold and hence, those samples were correctly classified as "A". Based on this observations, it can be concluded that using context-based feature along with physiological features can perform better than that using only physiological feature. This reduces false positives and false negatives and provides a robust model for distinguishing between anxious and non-anxious states.

5.3.4 Real-Time Anxiety Detection

To analyze the feasibility of the proposed method in detecting anxiety in real-time, the end-to-end process of anxiety detection is simulated for one minute. The simulation is performed on a computer with 16 GB RAM, and 3.3 GHz processor. Python 3.7 was used to simulate the results and the latency has been calculated using the time module of python. The end-to-end process consists of loading the data, signal processing, feature extraction, and prediction (Figure 5.1). The trained models (Table 5.4) are imported as

a .pkl file. Table 5.6 shows the simulation result for all the trained models (EDA, EDA+C, BVP, and BVP+C) in terms of latency, and the size of the trained model in memory.

Table 5.6 shows the memory of the trained models and the latency in seconds during each phase of the end-to-end processing. From Table 5.6, we can observe that the size of all the trained models is just in the order of a few hundred KB (KiloByte). Further, the latency during each phase is less than a second. The context-based models also performed almost similar to non-context-based models in terms of latency while occupying less in-device memory.

5.4 Conclusion

In this chapter, a hybrid physiological-context-based machine learning model for detecting anxiety in older adults is proposed. The proposed system uses two physiological sensors, EDA and PPG for recording EDA and BVP signals respectively during a TSST protocol. Anxiety level ground truth was obtained from the STAI questionnaire Features from EDA and BVP signals were used to evaluate three machine learning models, random forest, logistic regression, and support vector machine in the training phase. The best performing random forest classifier was optimized in the training phase using an exhaustive grid search technique for evaluation on the test data. Results on the test data showed that the optimized context-based version of EDA

and BVP models outperformed the non-context version of EDA and BVP models. The machine learning models that have used the context feature along with physiological features achieved higher F1-score for both the anxious and not-anxious states. Further, the feasibility of the proposed system for real-time anxiety detection in terms of latency, memory, and power is discussed.

In the existing literature, ECG and EEG signals are most commonly used for anxiety detection as these signals tend to be the most informative about anxiety levels. However, the acquisition setup required for recording these signals are usually complex and might be not be comfortable for daily use. Our proposed method for anxiety detection is based on simple wearable sensors such as EDA and PPG. These sensors can be easily integrated in low-cost consumer electronic devices such as a smart wristband. This will allow for continuous monitoring of anxiety levels with minimum obtrusiveness and will add to the comfort of the user. Further, our approach uses context feature in combination with physiological features instead of integrating features from other physiological signals. This is expected to minimize the design challenge faced when recording and processing information from multiple sensors in real-time. The current state of the proposed work on anxiety detection is suitable for controlled real-world settings where the activities (context) are more or less well defined. For example, a nursing home setting, or a training simulator could be some example use case for the proposed work.

In the future, combination of more fine-grained context features could be studied along with the physiological features to better estimate the anxiety level of a person. This could be especially useful for detecting anxiety under different activity such as walking, running etc. Investigating ways to seamlessly incorporate different context features from environmental sensors or activity sensors with the physiological features for anxiety detection in uncontrolled real-world setting could be another future research direction.

Chapter 6

Blood Pressure Estimation

In this chapter, we will present the work on blood pressure estimation using PPG based sensor. In this work we have proposed a computational framework for continuous blood pressure estimation using Photoplethysmogram (PPG) signal. The proposed framework is evaluated on the publicly available MIMIC Database. The database contains raw PPG data for different users and also the Arterial Blood Pressure (ABP) for calculating the systolic and diastolic blood pressure. We hypothesize that features extracted from the PPG signal and its derivatives can be used to estimate blood pressure values with significantly good accuracy. In this work, we have presented the validation of a single sensor and single probe PPG based computational framework in estimating blood pressure.

The rest of the chapter is organized as follows. Section 6.1 discusses the proposed framework. Section 6.2 presents the results and Section 6.3

concludes the chapter.

6.1 Proposed Computational Framework

This section presents an overview of the proposed framework (Figure 6.1)

The implementation of the proposed framework consists of following steps: (i) Data Extraction, (ii) Data Preprocessing, (iii) Feature Extraction, (iv) Feature Selection and (v) Predictive Modeling

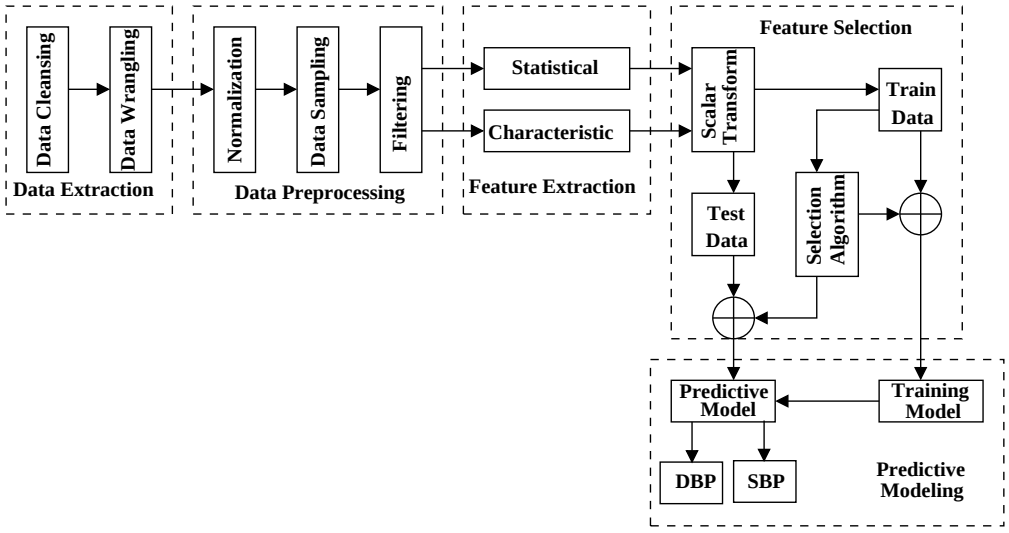


Figure 6.1: Overview of the proposed framework (© 2020 IEEE).

6.1.1 Data Extraction

In this work, we have used the MIMIC database to build and validate our model. The MIMIC database is a collection of multi parameter recordings of ICU patients [111]. The dataset was downloaded from PhysioBank database

which contains 72 complete records with varying length and sampling rate of 500 Hz. We have selected 20 records for our analysis. Two main selection criteria were used while selecting subjects for this study. First, only those subjects whose records included both PPG signal streams and Arterial Blood Pressure (ABP) signals were included. Secondly, out of the subjects whose PPG data were available, those subjects where the missing data consisted of more than 10% of the length of the data were filtered out. Also, the missing values of the PPG signal were replaced by the value 0. We have downloaded both short-term and long-term data for analysis.

For short-term analysis, we have downloaded 30 minutes of data from each record and for long-term analysis, we have downloaded 3 hours of data from each record. The three hours for each record were selected randomly one hour of data towards the beginning, middle and end of the entire record. The combined dataset consists of PPG signal and ABP signal. The ABP data is used for calculating the ground truth. The data preprocessing step is explained in detail in the next section.

6.1.2 Data Preprocessing

In this section, we will discuss the data preprocessing algorithm used in our blood pressure estimation model. The preprocessing algorithm consists of two parts: (i) Extracting PPG signal for different users and performing calibration (ii) Aggregating PPG signal of all subjects and performing noise removal. The proposed preprocessing algorithm is presented in Algorithm 1.

Algorithm 2: Proposed Data Preprocessing Algorithm (© 2020 IEEE).

Input : Raw PPG signal stream
Output: Processed PPG signal stream
1 Filter Order=N;
2 Cutoff Frequency=W_n;
3 **for** *all subjects* **do**
4 $\quad ppg\_subject \leftarrow extractPPGstreamforsubject$;
5 $\quad ppg\_mean \leftarrow mean(ppg\_subject)$;
6 $\quad ppg\_subject \leftarrow ppg\_subject - ppg\_mean$;
7 **end**
8 $df\_pleth \leftarrow concatenate(ppg\_subject)$;
9 $df\_pleth \leftarrow detrend(df\_pleth)$;
10 $df\_npleth \leftarrow normalize(df\_pleth)$;
11 $df\_modpleth \leftarrow df\_npleth\%mean(df\_npleth)$;
12 $df\_npleth \leftarrow df\_npleth - df\_modpleth$;
13 $df\_pleth\_filtered \leftarrow butterworth(df\_npleth, N, W_n)$;

The data proprocessing algorithm proposed, takes as input RAW PPG signal stream and generates processed PPG signal stream. The first and second line of the algorithm is the parameter initialization of the filter order and the cut off frequency represented by N and W_n respectively. In line 4, the algorithm extracts PPG signal from individual subjects. In line 5, the mean of the individual subject is calculated and in line 6, the mean is subtracted from the raw PPG value in order to remove the static component associated with the PPG signal ([112]). In line 8, the PPG signal stream obtained from the last step from different subjects is concatenated into a single frame.

After that, the PPG signal stream is further processed to remove the linear or static components that might still be associated with the PPG

signal stream (line 9). In line 10, the PPG signal is normalized, so that the data remains in the range [0,1]. This step removes any outliers which are likely to be noise. An additional processing step in the algorithm in line 11 and 12 is done to ensure further noise removal and proper curve fitting during regression analysis. The idea is to represent each sample point as a factor of the mean. For example, $x(t) = n.mean(x)$. where x is the signal stream and $x(t)$ is the value of the signal at any instant of time t and n is an integer. This step is done by calculating the residue when the sample point is divided by the mean of the signal by using the modulus operation and then subtracting the residue from the signal stream. The last step of the preprocessing stage is filtering of low frequency components. For, this reason, butterworth filter is selected to filter out any signal lower than 20 Hz frequency that is the cut off frequency. In the following section, we will discuss the feature extraction process of the proposed method.

6.1.3 Feature Extraction

Two types of features, statistical and characteristic features were extracted from the PPG signal. Features are extracted mainly from six signal streams one of which is the PPG signal stream and the rest five are derived from PPG signal stream. The five signal streams are: (i) first derivative of PPG signal, (ii) second derivative of PPG signal, (iii) frequency spectrum of PPG signal, (iv) first derivative of the frequency spectrum and (v) second derivative of frequency spectrum.

Statistical Features

By extracting statistical features over a sample of the PPG signal, we aim to obtain obtain its morphological information and map it to the target variable. This information is obtained by extracting these following measures:

- Measure of central tendency: Mean, Median, grouped median, harmonic mean, high median and low median.
- Measure of spread: Maximum, minimum, variance, standard deviation, root mean square, population variance and population standard deviation.
- Measure of shape: Kurtosis and skewness

For the original PPG signal and its frequency spectrum, we have extracted all the 15 statistical measure. However, for the derivatives of PPG signal and its frequency spectrum, we have excluded the grouped median, population variance and population standard deviation. Hence, a total of 78 statistical features extracted from all of the signal streams.

Characteristic Features

A total of 28 characteristic features were extracted. Out of which 16 features were extracted from PPG signal and 12 features from the power spectrum of the PPG signal. The 12 feature extracted from the power spectrum are mean, maximum, minimum, root mean square, kurtosis and skewness of the

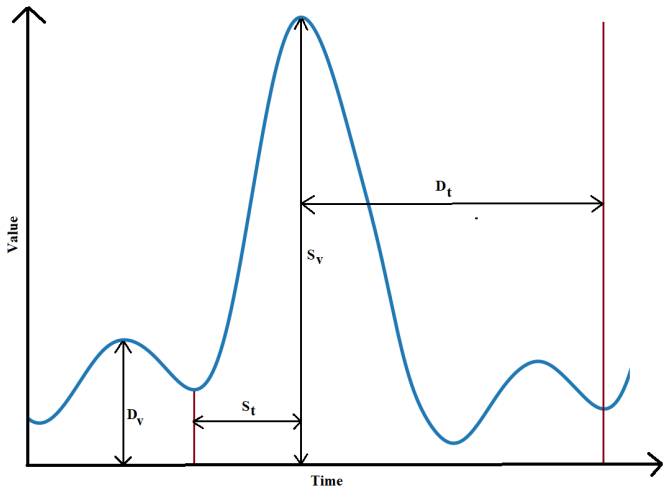


Figure 6.2: PPG signal and components (© 2020 IEEE).

peak amplitude and frequency. The 16 features extracted from PPG signal is described in Table 6.1 and represented in Figure 6.2.

Table 6.1: Characteristic feature description of the PPG signal (© 2020 IEEE).

Feature Name	**Feature Description**
S_t	Systolic time
D_t	Diastolic time
S_v	Systolic value
D_v	Diastolic value
dSDv	$\lvert S_v - D_v \rvert$
dSVt	$\lvert S_t - D_t \rvert$
rSDv	S_v / D_v
rSDt	S_t / D_t
ReSDv	$\lvert S_v - (S_v \mod D_v) \rvert$
ReSDt	$\lvert S_t - (S_t \mod D_t) \rvert$
ReDSv	$\lvert D_v - (D_v \mod S_v) \rvert$
ReDSt	$\lvert D_t - (D_t \mod S_t) \rvert$
mSDt	$S_t \mod D_t$
mDSt	$D_t \mod S_t$
mSDv	$S_v \mod D_v$
mDSv	$D_v \mod S_v$

The incorporation of the characteristic features from PPG signal into the feature set is done with a view to generating further information about the relation of the blood pressure values apart from the information obtained from only the statistical features of the PPG signal and its derivatives. Systolic time (S_t) and diastolic time (D_t) are validated in the literature to be important parameters in co-relating with the blood pressure values [113]. Apart from the systolic and diastolic time, we have calculated few other features like the difference, ratio and residue of between the systolic and diastolic time and values.

The reason for computing such parameters is important because the systolic and diastolic time alone just provides us with the absolute information of the contraction and relaxation of heart. However, computing the above mentioned features provides us with relative information of the systolic and diastolic time thereby establishing a relation between the two parameters. For instance, the difference, ratio provides us with how fast or slow the systolic time differs from that of diastolic time and by what factor.

A total of 106 features were computed from the PPG signal and its derivatives out of which there are 28 characteristics feature and 78 statistical features. The 28 characteristic feature consists of 16 purely characteristic feature as described in Table 6.1 and 12 features computed from the power spectrum which includes important statistical computation of the peak amplitude and frequency of the signal.

Although the features extracted from both the data sets are same, the

sampling window is not the same. For short-term data, features were extracted by taking a window size of 2 sec and overlap of 1 sec, For long-term data, features were extracted by taking a window size of 10 sec and overlap of 5 sec.

In the next section, we will discuss the feature selection module of the framework.

6.1.4 Feature Selection

The extracted feature from the feature extraction phase is forwarded to the feature selection unit. The feature set is first scalar transformed to map features from higher dimension to a lower dimension. This helps in identifying features with high multicollinearity. Two or more features with high multicollinearity gets transformed to similar features in a lower dimension and hence can be easily detected through testing of occurrence of duplicate features and can be removed. After scalar transformed, the feature set is split randomly into two sets: training set (70%) and test set (30 %). The train data which contains the training feature set and the training target data is fed into the selection algorithm (Algorithm 2).

The feature selection algorithm takes the train data as input and outputs the index of the selected feature. Line 1 of the algorithm denotes the maximum number of features that should be selected. Line 2 and 3 denotes the assignment of training feature vector and training target vector to X_train and Y_train respectively. In line 3, score for each feature is calculated using

Algorithm 3: Feature Selection Algorithm (© 2020 IEEE).

Input : Train Data
Output: Index of selected features

1 Maximum no. of features=K_{max};
2 $X\_train \leftarrow$ training feature vector ;
3 $Y\_train \leftarrow$ training target data;
4 $score \leftarrow mutual\_info\_regression(X\_train, Y\_train)$;
5 $sorted\_score \leftarrow sort(score)$;
6 $sorted\_score \leftarrow drop\_duplicates(score)$;
7 **for** i *in* $(0, K\_max)$ **do**
8 $\quad index\_select\_feature \leftarrow indexof(sorted\_score[i])$;
9 **end**
10 return $index\_select\_feature$

the $mutual\_info\_regression$ model from scikit learn [97]. This class returns an array containing the score of each feature based on mutual information between the feature vector and the target data.In line 5, the score is sorted and the duplicate scores are removed in line 7. Then from line 7 to line 9, index of top K_{max} features are extracted and returned in line 10.

6.1.5 Predictive Modeling

The continuous blood pressure prediction model estimates BP values from the PPG signal in some time window. We have implemented three kinds of regression model to train and test our data to estimate blood pressure values using the feature set generated. The models implemented are: (i) Decision Tree Regressor. (ii) Multi-layer Perceptron Regressor and (iii) Adaboost Regressor. Before the feature set is fed into the regression model, the feature

set is standardized using scalar transform ([97]).

6.2 Results

The training set is used to train the machine learning algorithms discussed above and the performance of the ML algorithms in estimating continuous blood pressure is discussed in this section. The ML algorithms are trained and tested on both short-term and long-term data.

We will evaluate the performance of continuous blood pressure predictive model using two metrics: Mean Absolute Error (MAE) and Standard Error (SE).

The analysis of the result is done in two phases:

- *Feature salience analysis*: In this analysis, we will observe the prediction accuracy in terms of mean absolute error and standard deviation for top 80 features. This analysis will be done on both short-term and long-term data.

- *Evaluation of model performance based on selected features*: In this analysis, we will estimate the performance of the predictive model using the selected combination of features.

6.2.1 Feature Salience Analysis

The short-term and long-term data are trained and tested with the AdaBoost regressor with decision tree as the base estimator with maximum depth of 30. Figure 6.3 shows the variation of mean absolute error for the 80 feature combination for systolic and diastolic prediction. Figure 6.4 shows the variation of standard deviation for the 80 feature combination for systolic and diastolic prediction.

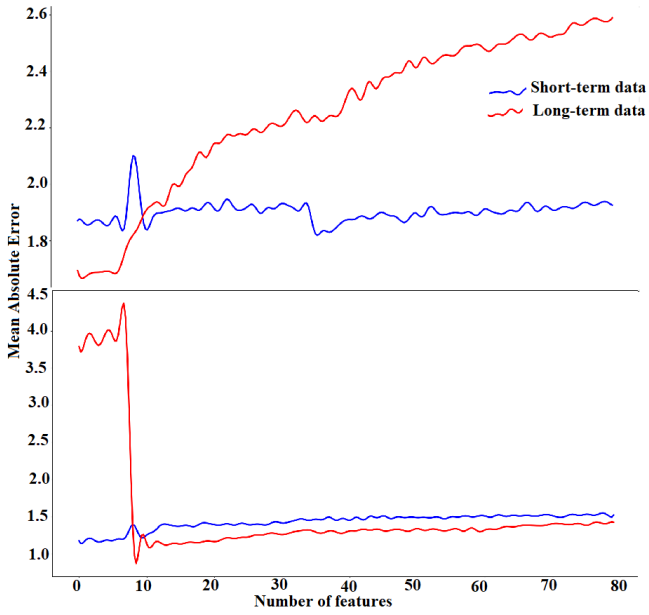


Figure 6.3: Plot of mean absolute error and number of features. (Top systolic and bottom diastolic) (© 2020 IEEE).

The following observations can be made by analyzing Figure 6.3 and Figure 6.4.

- MAE and SD of short-term data for both systolic and diastolic predic-

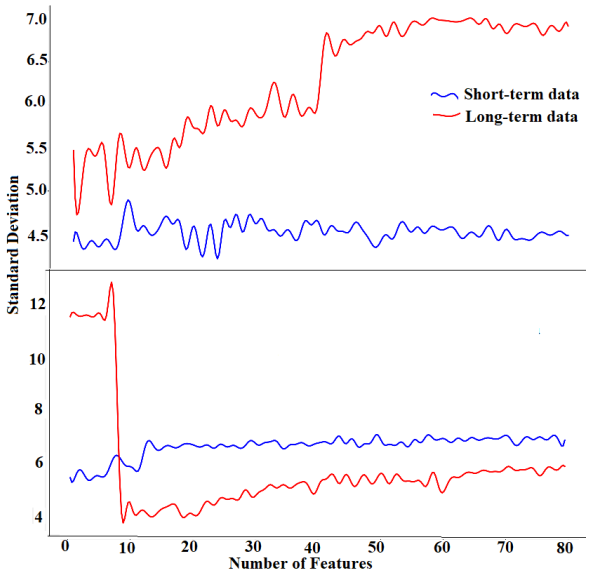


Figure 6.4: Plot of standard deviation and number of features. (Top systolic and bottom diastolic) (© 2020 IEEE).

tion remains more or less constant with the number of features.

- MAE and SD of long term data for systolic prediction follows an increasing trend with increase in number of features. However, for diastolic prediction, the MAE and SD both remains more or less constant with number of features after an initial abrupt drop.

- For diastolic prediction with long-term analysis, the MAE and SD optimizes when the the number of features is 10.

From the above observations, we can conclude that to optimize the systolic and diastolic prediction, we need to have atleast 10 features.

To understand which features would contribute best to the systolic and diastolic prediction, we identified the top 20 important features ranked by the

selection algorithm. Out of the top 20 features, the top 8 features were the statistical measure of the spectral characteristic of the PPG signal. These features are the maximum, minimum, skewness, kurtosis, mean and root mean square of the frequency and maximum and minimum of peak amplitude. The rest 12 features were strictly statistical measures from PPG signal and its derivatives. All of the 12 statistical feature were measure of spread, that is standard deviation, variance, maximum and minimum. Statistical features from fourier transform and its derivatives were not found to be important by the selection algorithm.

It is interesting to note that maximum and minimum of PPG and its first and second derivative were all among the 12 statistical feature and the 8 characteristic features also contained maximum and minimum of frequency and peaks. Hence, we have used the maximum and minimum of frequency, peaks, PPG signal, PPG signal first derivative and PPG signal second derivative to build a feature set of 10 features.

6.2.2 Evaluation of Model Performance Based on Selected Feature

In this section, we will discuss the results of the performance of the four regression models based on the selected 10 features. The performance is evaluated using mean absolute error , standard deviation and percentage of error less than 5 mmHg. The results for both short-term and long-term

analysis are presented in Table 6.2 for systolic and Table 6.3 for diastolic prediction

Table 6.2: Performance analysis for SBP prediction (© 2020 IEEE).

	MLP Regressor		**DT Regressor**		**AdaBoost(DT)**		**AdaBoost(MLP)**	
	Short-term	*Long-term*	*Short-term*	*Long-term*	*Short-term*	*Long-term*	*Short-term*	*Long-term*
MAE	11.29	16.20	2.34	2.56	1.69	2.07	22.05	19.59
SD	12.35	15.51	5.42	6.00	4.52	5.97	14.85	15.66
% MAE	36	23	87	86	93	91	13.8	17

Table 6.3: Performance analysis for DBP prediction (© 2020 IEEE).

	MLP Regressor		**DT Regressor**		**AdaBoost(DT)**		**AdaBoost(MLP)**	
	Short-term	*Long-term*	*Short-term*	*Long-term*	*Short-term*	*Long-term*	*Short-term*	*Long-term*
MAE	5.23	6.6	1.69	1.55	1.32	1.15	11.01	12.77
SD	9.12	8.51	6.74	6.41	6.40	4.05	13.89	16.38
% Error	67	54	94	93	97	96	45	39

Performance of Predictive Model on Short-term Analysis

From Table 6.2 and Table 6.3, it can be observed that AdaBoost Regressor with Decision Tree as the base estimator performs the best both for SBP and DBP prediction with an MAE of 1.69 and SD of 4.52 for SBP and an MAE of 1.32 and SD of 6.40.

Performance of Predictive Model on Long-term Analysis

From Table 6.2 and Table 6.3, it can be observed that AdaBoost Regressor with Decision Tree as the base estimator performs the best for both SBP and DBP prediction with an MAE of 2.07 and SE of 5.97 for SBP prediction and an MAE of 1.15 and SE of 4.05 for DBP prediction.

6.2.3 Comparison with Existing Works

The results obtained in this work is compared with the existing works. Table 6.4 shows the comparison between the proposed work and the existing works in terms of mean absolute error and standard deviation. As all the work were validated on different dataset, we have also also listed the dataset used for each of these works. From the table, we can observe that using PPG alone performs significantly well in terms of MAE and SE.

Table 6.4: Comparison Table (© 2020 IEEE).

Work	SBP		DBP		Data	
	MAE	SD	MAE	SD	Signals	Data set
[89]	9.58	NA	5.26	NA	PPG, ECG	Experimental
[55]	6.22	9.44	3.97	5.15	PCG, PPG	Experimental
[55]	4.71	6.15	4.44	5.36	ECG, PPG	Experimental
[90]	9	8.58	7	5.81	BCG, PPG	Experimental
[114]	8.21	5.45	4.31	3.52	ECG, PPG	MIMIC 2
This work	**2.07**	**5.97**	**1.15**	**4.05**	**PPG**	**MIMIC**

6.3 Conclusion

The work presented in this chapter attempts to estimate systolic and diastolic blood pressure. The raw PPG data was obtained across several patients from the MIMIC database. A feature selection algorithm was proposed to select important features that contribute the most towards optimizing systolic and diastolic prediction error. Final feature set which consisted of only 10 features was selected for further evaluation. Analysis was carried on both short-term and long-term data. Different machine learning algorithm are implemented

to estimate blood pressure from the feature set. The performance results showed that the AdaBoost Regressor with decision tree as the base estimator performed best in estimating blood pressure values for both short-term and long-term data. Results indicated that the proposed model based on single signal (PPG) and single probe approach is able to estimate systolic and diastolic blood pressure with significantly high accuracy. A single sensor, single probe measurement of PPG signal is ideal for wearable devices adding convenience to the user. This approach has potential to be a good fit for smart home environment where continuous monitoring of blood pressure in an unobtrusive manner will facilitate long term in home care.

Chapter 7

Conclusion and Future Directions

Monitoring and managing stress and the accompanying anxiety and high blood pressure regularly can prevent long-term damages and prevent these conditions from becoming chronic. Historical health records generated from long-term monitoring of stress, anxiety, and blood pressure can reveal interesting behavior and living pattern which could also be useful for the diagnosis of related health problems. In this dissertation, we have proposed technological solutions for detecting stress, anxiety and estimating blood pressure values by monitoring physiological signals. We have used machine learning models for correlating the features extracted from the physiological signals with the ground truth for developing the predictive models. We have considered those physiological signals that can be easily recorded with minimum

obtrusiveness such as fingertip-based sensors or wristband-based sensors.

Based on the experimental data collection of 50 healthy older adults we have designed, developed, and validated stress detection models for a fingertip-based wearable device and a wristband-based wearable device. The fingertip-based wearable devices used EDA and BVP signals from EDA and PPG fingertip sensors. A deep-learning-based LSTM classifier was proposed for stress level classification along with benchmark machine learning classifiers. Results showed that the proposed LSTM performed slightly better than the traditional machine learning algorithms. The wristband-based wearable devices used EDA, PPG, and ST sensors for extracting EDA, BVP, IBI, and ST signals for stress detection. Results have shown that integrating features from multiple signals helps improve the classification performance significantly by reducing the number of false positives. A voice-based interactive system for stress level query has been prototyped and simulated for integrating the proposed stress detection system in a consumer end device.

An anxiety detection model that uses a single wearable sensor and a context feature has been proposed. The proposed model evaluated the combination of EDA and BVP features with the experimental context feature individually for classifying between anxious and not-anxious states. Results on the test data showed that the optimized context-based version of EDA and BVP models outperformed the non-context version of EDA and BVP models. The machine learning models that have used the context feature along with physiological features achieved higher F1-score for both the anx-

ious and not-anxious states. Finally, a computational framework for blood pressure estimation using a single PPG sensor is proposed. The proposed framework is validated using a freely available online MIMIC database.

The research explored in this dissertation is an important step towards realizing a ubiquitous health monitoring system with extensive diagnostic capabilities. Implementing the proposed designs for stress, anxiety, and blood pressure on low-power microcontroller devices by optimizing the computations while maintaining the performance could be an interesting future research direction. Another interesting research direction could be the use of emerging computing paradigms such as quantum annealing for optimizing the predictive model during the training phase.

Bibliography

[1] J. Bakker, M. Pechenizkiy, and N. Sidorova, "What's your current stress level? detection of stress patterns from gsr sensor data," in *2011 IEEE 11th international conference on data mining workshops*, pp. 573–580, IEEE, 2011.

[2] P. Warr and K. Nielsen, "Wellbeing and work performance," *Handbook of well-being. Salt Lake City, UT: DEF Publishers*, 2018.

[3] V. Koubova and A. A. Buchko, "Life-work balance: Emotional intelligence as a crucial component of achieving both personal life and work performance," *Management Research Review*, 2013.

[4] G. P. Chrousos, "Stress and disorders of the stress system," *Nature reviews endocrinology*, vol. 5, no. 7, p. 374, 2009.

[5] D. of Health, D. Human Services, Washington, H. P. . (Group), and U. S. G. P. Office, *Healthy people 2010: Understanding and improving health*. US Department of Health and Human Services, 2000.

[6] "The American Institute of Stress, `https://www.stress.org/`, year=2020, Accessed: 2021-05-25."

[7] "American Psychological Association, `https://www.apa.org/news/press/releases/stress/2020/report-october#:~:text=nearly%201%20in%205%20adults,and%20older%20adults%20(8%25).ttps://www.stress.org/`, year=2020, Accessed: 2021-05-25."

[8] K. T. Larkin, *Stress and hypertension: examining the relation between psychological stress and high blood pressure*. Yale university press, 2008.

[9] C. Grillon, R. Duncko, M. F. Covington, L. Kopperman, and M. A. Kling, "Acute stress potentiates anxiety in humans," *Biological psychiatry*, vol. 62, no. 10, pp. 1183–1186, 2007.

[10] "American Psychological Association, `https://www.apa.org/news/press/releases/stress/2017/state-nation.pdf`, year=2017, Accessed: 2021-05-27."

[11] C. Bryant, H. Jackson, and D. Ames, "The prevalence of anxiety in older adults: methodological issues and a review of the literature," *J Affect Disord*, vol. 109, pp. 233–250, Aug 2008.

[12] A. T. Beekman, E. de Beurs, A. J. van Balkom, D. J. Deeg, R. van Dyck, and W. van Tilburg, "Anxiety and depression in later life: Co-occurrence and communality of risk factors," *Am J Psychiatry*, vol. 157, pp. 89–95, Jan 2000.

[13] I. Maatouk, W. Herzog, F. Böhlen, R. Quinzler, B. Löwe, K. U. Saum, H. Brenner, and B. Wild, "Association of hypertension with depression and generalized anxiety symptoms in a large population-based sample of older adults," *J Hypertens*, vol. 34, pp. 1711–1720, 09 2016.

[14] S. Cohen, R. C. Kessler, L. U. Gordon, *et al.*, "Strategies for measuring stress in studies of psychiatric and physical disorders," *Measuring stress: A guide for health and social scientists*, pp. 3–26, 1995.

[15] B. S. McEwen, "Stress, adaptation, and disease: Allostasis and allostatic load," *Annals of the New York academy of sciences*, vol. 840, no. 1, pp. 33–44, 1998.

[16] T. Fulop, A. Larbi, J. M. Witkowski, J. McElhaney, M. Loeb, A. Mitnitski, and G. Pawelec, "Aging, frailty and age-related diseases," *Biogerontology*, vol. 11, no. 5, pp. 547–563, 2010.

[17] G. E. Miller, N. Rohleder, C. Stetler, and C. Kirschbaum, "Clinical depression and regulation of the inflammatory response during acute stress," *Psychosomatic medicine*, vol. 67, no. 5, pp. 679–687, 2005.

[18] J. E. Graham, L. M. Christian, and J. K. Kiecolt-Glaser, "Stress, age, and immune function: toward a lifespan approach," *Journal of behavioral medicine*, vol. 29, no. 4, pp. 389–400, 2006.

[19] G. Gerra, D. Monti, A. E. Panerai, P. Sacerdote, R. Anderlini, P. Avanzini, A. Zaimovic, F. Brambilla, and C. Franceschi, "Long-term immune-endocrine effects of bereavement: relationships with anxiety levels and mood," *Psychiatry research*, vol. 121, no. 2, pp. 145–158, 2003.

[20] J. K. Kiecolt-Glaser and R. Glaser, "Stress and immune function in humans," in *Psychoneuroimmunology*, pp. 849–867, Elsevier, 1991.

[21] S. C. Segerstrom and G. E. Miller, "Psychological stress and the human immune system: a meta-analytic study of 30 years of inquiry.," *Psychological bulletin*, vol. 130, no. 4, p. 601, 2004.

[22] N. Daviu, M. R. Bruchas, B. Moghaddam, C. Sandi, and A. Beyeler, "Neurobiological links between stress and anxiety," *Neurobiology of stress*, vol. 11, p. 100191, 2019.

[23] Y. Takagi, Y. Sakai, Y. Abe, S. Nishida, B. J. Harrison, I. Martínez-Zalacaín, C. Soriano-Mas, J. Narumoto, and S. C. Tanaka, "A common brain network among state, trait, and pathological anxiety from whole-brain functional connectivity," *Neuroimage*, vol. 172, pp. 506–516, 2018.

[24] J. Fawcett, "The detection and consequences of anxiety in clinical depression," *Journal of Clinical Psychiatry*, vol. 58, no. 8, p. 35, 1997.

[25] C. A. Emdin, A. Odutayo, C. X. Wong, J. Tran, A. J. Hsiao, and B. H. Hunn, "Meta-analysis of anxiety as a risk factor for cardiovascular disease," *The American journal of cardiology*, vol. 118, no. 4, pp. 511–519, 2016.

[26] K. E. Vytal, B. R. Cornwell, N. E. Arkin, A. M. Letkiewicz, and C. Grillon, "The complex interaction between anxiety and cognition: insight from spatial and verbal working memory," *Frontiers in human neuroscience*, vol. 7, p. 93, 2013.

[27] G. N. Papadimitriou, M. Kerkhofs, C. Kempenaers, and J. Mendlewicz, "Eeg sleep studies in patients with generalized anxiety disorder," *Psychiatry research*, vol. 26, no. 2, pp. 183–190, 1988.

[28] A. J. Flint, "Anxiety and its disorders in late life: moving the field forward.," 2005.

[29] F. Edition *et al.*, "Diagnostic and statistical manual of mental disorders," *Am Psychiatric Assoc*, vol. 21, 2013.

[30] K. M. Mehta, E. M. Simonsick, B. W. Penninx, R. Schulz, S. M. Rubin, S. Satterfield, and K. Yaffe, "Prevalence and correlates of anxiety symptoms in well-functioning older adults: findings from the health aging and body composition study," *Journal of the American Geriatrics Society*, vol. 51, no. 4, pp. 499–504, 2003.

[31] H. P. Van Hout, A. T. Beekman, E. De Beurs, H. Comijs, H. Van Marwijk, M. De Haan, W. Van Tilburg, and D. J. Deeg, "Anxiety and the risk of death in older men and women," *The British Journal of Psychiatry*, vol. 185, no. 5, pp. 399–404, 2004.

[32] J. L. Wetherell, S. R. Thorp, T. L. Patterson, S. Golshan, D. V. Jeste, and M. Gatz, "Quality of life in geriatric generalized anxiety disorder: a preliminary investigation," *Journal of Psychiatric Research*, vol. 38, no. 3, pp. 305–312, 2004.

[33] T. M. Spruill, "Chronic psychosocial stress and hypertension," *Curr Hypertens Rep*, vol. 12, pp. 10–16, Feb 2010.

[34] C. Nordqvist, "What's to know about high blood pressure," 2017.

[35] J. E. Schwartz, T. G. Pickering, and P. A. Landsbergis, "Work-related stress and blood pressure: current theoretical models and considerations from a behavioral medicine perspective," *J Occup Health Psychol*, vol. 1, pp. 287–310, Jul 1996.

[36] T. G. Vrijkotte, L. J. van Doornen, and E. J. de Geus, "Effects of work stress on ambulatory blood pressure, heart rate, and heart rate variability," *Hypertension*, vol. 35, pp. 880–886, Apr 2000.

[37] A. Steptoe, J. Siegrist, C. Kirschbaum, and M. Marmot, "Effort-reward imbalance, overcommitment, and measures of cortisol and blood pressure over the working day," *Psychosom Med*, vol. 66, no. 3, pp. 323–329, 2004.

[38] J. B. Byrd and R. D. Brook, "Anxiety in the "age of hypertension"," *Current hypertension reports*, vol. 16, no. 10, pp. 1–7, 2014.

[39] B. Löwe, K. Gräfe, C. Ufer, K. Kroenke, E. Grünig, W. Herzog, and M. M. Borst, "Anxiety and depression in patients with pulmonary hypertension," *Psychosomatic medicine*, vol. 66, no. 6, pp. 831–836, 2004.

[40] H. Thapliyal, R. Kumar Nath, and S. P. Mohanty, "Smart home environment for mild cognitive impairment population: Solutions to improve care and quality of life," *IEEE Consumer Electronics Magazine*, vol. 7, no. 1, pp. 68–76, 2018.

[41] R. K. Nath, H. Thapliyal, A. Caban-Holt, and S. P. Mohanty, "Machine learning based solutions for real-time stress monitoring," *IEEE Consumer Electronics Magazine*, vol. 9, no. 5, pp. 34–41, 2020.

[42] S. Greene, H. Thapliyal, and D. Carpenter, "Iot-based fall detection for smart home environments," in *2016 IEEE International Symposium on Nanoelectronic and Information Systems (iNIS)*, pp. 23–28, 2016.

[43] R. K. Nath, R. Bajpai, and H. Thapliyal, "Iot based indoor location detection system for smart home environment," in *2018 IEEE International Conference on Consumer Electronics (ICCE)*, pp. 1–3, 2018.

[44] R. K. Nath and H. Thapliyal, "Wearable health monitoring system for older adults in a smart home environment," in *IEEE Computer Society Annual Symposium on VLSI, 2021 (Accepted for Publication)*.

[45] H. Thapliyal, V. Khalus, and C. Labrado, "Stress detection and management: A survey of wearable smart health devices," *IEEE Consumer Electronics Magazine*, vol. 6, no. 4, pp. 64–69, 2017.

[46] Z. Li, X. Wu, X. Xu, H. Wang, Z. Guo, Z. Zhan, and L. Yao, "The recognition of multiple anxiety levels based on electroencephalograph," *IEEE Transactions on Affective Computing*, 2019.

[47] K. Song, K.-y. Chung, and J.-H. Chang, "Cuffless deep learning-based blood pressure estimation for smart wristwatches," *IEEE Transactions on Instrumentation and Measurement*, vol. 69, no. 7, pp. 4292–4302, 2020.

[48] S. Greene, H. Thapliyal, and A. Caban-Holt, "A survey of affective computing for stress detection: Evaluating technologies in stress detection for better health," *IEEE Consumer Electronics Magazine*, vol. 5, no. 4, pp. 44–56, 2016.

[49] 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*, pp. 1–13, 2021.

[50] R. K. Nath, H. Thapliyal, and A. Caban-Holt, "Validating physiological stress detection model using cortisol as stress bio marker," in *2020 IEEE International Conference on Consumer Electronics (ICCE)*, pp. 1–5, 2020.

[51] R. K. Nath and H. Thapliyal, "Smart wristband-based stress detection framework for older adults with cortisol as stress biomarker," *IEEE Transactions on Consumer Electronics*, vol. 67, no. 1, pp. 30–39, 2021.

[52] R. K. Nath and H. Thapliyal, "Machine learning based anxiety detection in older adults using wristband sensors and context feature," *Springer Nature Computer Science (Accepted for Publication)*, 2021.

[53] R. K. Nath and H. Thapliyal, "Ppg based continuous blood pressure monitoring framework for smart home environment," in *2020 IEEE 6th World Forum on Internet of Things (WF-IoT)*, pp. 1–6, 2020.

[54] R. K. Nath, H. Thapliyal, and A. Caban-Holt, "Towards photoplethysmogram based non-invasive blood pressure classification," in *2018 IEEE International Symposium on Smart Electronic Systems (iSES) (Formerly iNiS)*, pp. 37–39, 2018.

[55] A. Esmaili, M. Kachuee, and M. Shabany, "Nonlinear cuffless blood pressure estimation of healthy subjects using pulse transit time and arrival time," *IEEE Transactions on Instrumentation and Measurement*, vol. 66, no. 12, pp. 3299–3308, 2017.

[56] S. Betti, R. M. Lova, E. Rovini, G. Acerbi, L. Santarelli, M. Cabiati, S. Del Ry, and F. Cavallo, "Evaluation of an integrated system of

wearable physiological sensors for stress monitoring in working environments by using biological markers," *IEEE Transactions on Biomedical Engineering*, vol. 65, no. 8, pp. 1748–1758, 2017.

[57] W. Wen, G. Liu, Z.-H. Mao, W. Huang, X. Zhang, H. Hu, J. Yang, and W. Jia, "Toward constructing a real-time social anxiety evaluation system: Exploring effective heart rate features," *IEEE Transactions on Affective Computing*, vol. 11, no. 1, pp. 100–110, 2018.

[58] E. Smets, W. De Raedt, and C. Van Hoof, "Into the wild: the challenges of physiological stress detection in laboratory and ambulatory settings," *IEEE journal of biomedical and health informatics*, vol. 23, no. 2, pp. 463–473, 2018.

[59] X. Zhang, J. Pan, J. Shen, Z. U. Din, J. Li, D. Lu, M. Wu, and B. Hu, "Fusing of electroencephalogram and eye movement with group sparse canonical correlation analysis for anxiety detection," *IEEE Transactions on Affective Computing*, 2020.

[60] F. Delmastro, F. Di Martino, and C. Dolciotti, "Cognitive training and stress detection in mci frail older people through wearable sensors and machine learning," *IEEE Access*, vol. 8, pp. 65573–65590, 2020.

[61] U. R. Acharya, S. V. Sree, P. C. A. Ang, R. Yanti, and J. S. Suri, "Application of non-linear and wavelet based features for the automated identification of epileptic eeg signals," *International journal of neural systems*, vol. 22, no. 02, p. 1250002, 2012.

[62] E. Ferreira, D. Ferreira, S. Kim, P. Siirtola, J. Röning, J. F. Forlizzi, and A. K. Dey, "Assessing real-time cognitive load based on psycho-physiological measures for younger and older adults," in *2014 IEEE Symposium on Computational Intelligence, Cognitive Algorithms, Mind, and Brain (CCMB)*, pp. 39–48, IEEE, 2014.

[63] R. A. McFarland, "Relationship of skin temperature changes to the emotions accompanying music," *Biofeedback and Self-regulation*, vol. 10, no. 3, pp. 255–267, 1985.

[64] M.-h. Lee, G. Yang, H.-K. Lee, and S. Bang, "Development stress monitoring system based on personal digital assistant (pda)," in *The 26th*

Annual International Conference of the IEEE Engineering in Medicine and Biology Society, vol. 1, pp. 2364–2367, IEEE, 2004.

[65] P. Yuen, K. Hong, T. Chen, A. Tsitiridis, F. Kam, J. Jackman, D. James, M. Richardson, L. Williams, W. Oxford, *et al.*, "Emotional & physical stress detection and classification using thermal imaging technique," 2009.

[66] G. Giannakakis, D. Grigoriadis, K. Giannakaki, O. Simantiraki, A. Roniotis, and M. Tsiknakis, "Review on psychological stress detection using biosignals," *IEEE Transactions on Affective Computing*, 2019.

[67] S. Yoon, J. K. Sim, and Y.-H. Cho, "A flexible and wearable human stress monitoring patch," *Scientific reports*, vol. 6, no. 1, pp. 1–11, 2016.

[68] Y. Lee, B. Lee, and M. Lee, "Wearable sensor glove based on conducting fabric using electrodermal activity and pulse-wave sensors for e-health application," *Telemedicine and e-Health*, vol. 16, no. 2, pp. 209–217, 2010.

[69] R. R. Fletcher, K. Dobson, M. S. Goodwin, H. Eydgahi, O. Wilder-Smith, D. Fernholz, Y. Kuboyama, E. B. Hedman, M.-Z. Poh, and R. W. Picard, "icalm: Wearable sensor and network architecture for wirelessly communicating and logging autonomic activity," *IEEE transactions on information technology in biomedicine*, vol. 14, no. 2, pp. 215–223, 2010.

[70] M. Quazi, S. Mukhopadhyay, N. Suryadevara, and Y.-M. Huang, "Towards the smart sensors based human emotion recognition," in *2012 IEEE International Instrumentation and Measurement Technology Conference Proceedings*, pp. 2365–2370, IEEE, 2012.

[71] J. A. Healey and R. W. Picard, "Detecting stress during real-world driving tasks using physiological sensors," *IEEE Transactions on intelligent transportation systems*, vol. 6, no. 2, pp. 156–166, 2005.

[72] A. de Santos Sierra, C. S. Ávila, J. G. Casanova, and G. B. del Pozo, "A stress-detection system based on physiological signals and fuzzy logic," *IEEE Transactions on Industrial Electronics*, vol. 58, no. 10, pp. 4857–4865, 2011.

[73] C. Setz, B. Arnrich, J. Schumm, R. La Marca, G. Tröster, and U. Ehlert, "Discriminating stress from cognitive load using a wearable eda device," *IEEE Transactions on information technology in biomedicine*, vol. 14, no. 2, pp. 410–417, 2009.

[74] A. Puli and A. Kushki, "Toward automatic anxiety detection in autism: A real-time algorithm for detecting physiological arousal in the presence of motion," *IEEE Transactions on Biomedical Engineering*, vol. 67, no. 3, pp. 646–657, 2019.

[75] L. Mou, C. Zhou, P. Zhao, B. Nakisa, M. N. Rastgoo, R. Jain, and W. Gao, "Driver stress detection via multimodal fusion using attention-based cnn-lstm," *Expert Systems with Applications*, vol. 173, p. 114693, 2021.

[76] R. K. Nath, H. Thapliyal, and T. S. Humble, "A review of machine learning classification using quantum annealing for real-world applications," *Springer Nature Computer Science (Accepted for Publication)*, 2021.

[77] R. K. Nath, H. Thapliyal, and T. S. Humble, "Quantum annealing for automated feature selection in stress detection," in *IEEE Computer Society Annual Symposium on VLSI, 2021 (Accepted for Publication).*

[78] J. Clark, R. K. Nath, and H. Thapliyal, "Machine learning based prediction of future stress events in a driving scenario," in *2021, IEEE 7th World Forum on Internet of Things (Accepted for publication).*

[79] M. Sokolova and G. Lapalme, "A systematic analysis of performance measures for classification tasks," *Information processing & management*, vol. 45, no. 4, pp. 427–437, 2009.

[80] Wenhui Liao, Weihong Zhang, Zhiwei Zhu, and Qiang Ji, "A real-time human stress monitoring system using dynamic bayesian network," in *2005 IEEE Computer Society Conf. on Computer Vision and Pattern Recognition (CVPR'05) - Workshops*, pp. 70–70, 2005.

[81] J. J. Shaughnessy, E. B. Zechmeister, and J. S. Zechmeister, *Research methods in psychology.* McGraw-Hill, 2000.

[82] N. Charness, R. Best, and J. Evans, "Supportive home health care technology for older adults: Attitudes and implementation," *Gerontechnology: international journal on the fundamental aspects of technology to serve the ageing society*, vol. 15, no. 4, p. 233, 2016.

[83] B. Kikhia, T. G. Stavropoulos, S. Andreadis, N. Karvonen, I. Kompatsiaris, S. Sävenstedt, M. Pijl, and C. Melander, "Utilizing a wristband sensor to measure the stress level for people with dementia," *Sensors*, vol. 16, no. 12, p. 1989, 2016.

[84] M. Belk, D. Portugal, P. Germanakos, J. Quintas, E. Christodoulou, and G. S. Samaras, "A computer mouse for stress identification of older adults at work," vol. 1618.

[85] S.-M. Cheong, C. Bautista, and L. Ortiz, "Sensing physiological change and mental stress in older adults from hot weather," *IEEE Access*, vol. 8, pp. 70171–70181, 2020.

[86] L. Rachakonda, S. P. Mohanty, E. Kougianos, and P. Sundaravadivel, "Stress-lysis: A dnn-integrated edge device for stress level detection in the iomt," *IEEE Transactions on Consumer Electronics*, vol. 65, no. 4, pp. 474–483, 2019.

[87] Y. Zheng, T. C. H. Wong, B. H. K. Leung, and C. C. Y. Poon, "Unobtrusive and multimodal wearable sensing to quantify anxiety," *IEEE Sensors Journal*, vol. 16, no. 10, pp. 3689–3696, 2016.

[88] E. W. McGinnis, S. P. Anderau, J. Hruschak, R. D. Gurchiek, N. L. Lopez-Duran, K. Fitzgerald, K. L. Rosenblum, M. Muzik, and R. S. McGinnis, "Giving voice to vulnerable children: Machine learning analysis of speech detects anxiety and depression in early childhood," *IEEE Journal of Biomedical and Health Informatics*, vol. 23, no. 6, pp. 2294–2301, 2019.

[89] P. Nabeel, S. Karthik, J. Joseph, and M. Sivaprakasam, "Arterial blood pressure estimation from local pulse wave velocity using dual-element photoplethysmograph probe," *IEEE Transactions on Instrumentation and Measurement*, vol. 67, no. 6, pp. 1399–1408, 2018.

[90] C.-S. Kim, A. M. Carek, R. Mukkamala, O. T. Inan, and J.-O. Hahn, "Ballistocardiogram as proximal timing reference for pulse transit time

measurement: Potential for cuffless blood pressure monitoring," *IEEE Transactions on Biomedical Engineering*, vol. 62, no. 11, pp. 2657–2664, 2015.

[91] W.-H. Lin, D. Wu, C. Li, H. Zhang, and Y.-T. Zhang, "Comparison of heart rate variability from ppg with that from ecg," in *The International Conference on Health Informatics*, pp. 213–215, Springer, 2014.

[92] D. Bozovic, M. Racic, and N. Ivkovic, "Salivary cortisol levels as a biological marker of stress reaction," *Med Arch*, vol. 67, no. 5, pp. 374–377, 2013.

[93] M. Birkett, "The trier social stress test protocol for inducing psychological stress," *Journal of visualized experiments : JoVE*, 10 2011.

[94] S. Dickerson and M. Kemeny, "Acute stressors and cortisol responses: A theoretical integration and synthesis of laboratory research," *Psychological bulletin*, vol. 130, pp. 355–91, 06 2004.

[95] H. D. Critchley, "Electrodermal responses: what happens in the brain," *The Neuroscientist*, vol. 8, no. 2, pp. 132–142, 2002.

[96] S. Anders, M. Lotze, M. Erb, W. Grodd, and N. Birbaumer, "Brain activity underlying emotional valence and arousal: A response-related fmri study," *Human brain mapping*, vol. 23, no. 4, pp. 200–209, 2004.

[97] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay, "Scikit-learn: Machine learning in Python," *Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, 2011.

[98] S. Kulkarni, I. O'Farrell, M. Erasi, and M. Kochar, "Stress and hypertension.," *WMJ: official publication of the State Medical Society of Wisconsin*, vol. 97, no. 11, pp. 34–38, 1998.

[99] T. Bombardini, V. Gemignani, E. Bianchini, L. Venneri, C. Petersen, E. Pasanisi, L. Pratali, D. Alonso-Rodriguez, M. Pianelli, F. Faita, *et al.*, "Diastolic time–frequency relation in the stress echo lab: filling timing and flow at different heart rates," *Cardiovascular ultrasound*, vol. 6, no. 1, p. 15, 2008.

[100] F. Martini *et al.*, *Anatomy and Physiology'2007 Ed.* Rex Bookstore, Inc., 2006.

[101] V. C. Scanlon and T. Sanders, *Essentials of anatomy and physiology.* 2018.

[102] M. Yamaguchi and J. Sakakima, "Evaluation of driver stress in a motor-vehicle driving simulator using a biochemical marker," *Journal of international medical research*, vol. 35, no. 1, pp. 91–100, 2007.

[103] M. G. Kendall, "A new measure of rank correlation," *Biometrika*, vol. 30, no. 1/2, pp. 81–93, 1938.

[104] M. Fechir, T. Schlereth, T. Purat, S. Kritzmann, C. Geber, T. Eberle, M. Gamer, and F. Birklein, "Patterns of sympathetic responses induced by different stress tasks," *Open Neurol J*, vol. 2, pp. 25–31, 2008.

[105] T. C. Panagiotakopoulos, D. P. Lyras, M. Livaditis, K. N. Sgarbas, G. C. Anastassopoulos, and D. K. Lymberopoulos, "A contextual data mining approach toward assisting the treatment of anxiety disorders," *IEEE Transactions on Information Technology in Biomedicine*, vol. 14, no. 3, pp. 567–581, 2010.

[106] D. Adolph, L. Meister, and B. M. Pause, "Context counts! social anxiety modulates the processing of fearful faces in the context of chemosensory anxiety signals," *Front Hum Neurosci*, vol. 7, p. 283, 2013.

[107] C. D. Spielberger, "State-trait anxiety inventory," *The Corsini encyclopedia of psychology*, pp. 1–1, 2010.

[108] S. Nigussie, T. Belachew, and W. Wolancho, "Predictors of preoperative anxiety among surgical patients in jimma university specialized teaching hospital, south western ethiopia," *BMC surgery*, vol. 14, no. 1, pp. 1–10, 2014.

[109] K. S. Dalal, S. Chellam, and P. Toal, "Anaesthesia information booklet: Is it better than a pre-operative visit?," *Indian journal of anaesthesia*, vol. 59, no. 8, p. 511, 2015.

[110] O. Kayikcioglu, S. Bilgin, G. Seymenoglu, and A. Deveci, "State and trait anxiety scores of patients receiving intravitreal injections," *Biomedicine hub*, vol. 2, no. 2, pp. 1–5, 2017.

[111] A. Goldberger, L. Amaral, L. Glass, S. Havlin, J. M. Hausdorg, P. Ivanov, R. G. Mark, J. E. Mietus, G. B. Moody, C.-K. Peng, H. Stanley, and P. Physiobank, "Components of a new research resource for complex physiologic signals," *PhysioNet*, vol. 101, 01 2000.

[112] J. Sinex, "Pulse oximetry: Principles and limitations," *The American journal of emergency medicine*, vol. 17, pp. 59–67, 02 1999.

[113] J. Wikstrand, G. Berglund, L. Wilhelmsen, and I. Wallentin, "Value of systolic and diastolic time intervals. studies in normotensive and hypertensive 50-year-old men and in patients after myocardial infarction," *British heart journal*, vol. 40, pp. 256–67, 04 1978.

[114] M. Kachuee, M. M. Kiani, H. Mohammadzade, and M. Shabany, "Cuffless blood pressure estimation algorithms for continuous health-care monitoring," *IEEE Transactions on Biomedical Engineering*, vol. 64, pp. 859–869, April 2017.

Vita

Rajdeep Kumar Nath

Education

National Institute of Technology Durgapur
Bachelor of Technology in Computer Science and Engineering, May 2016

Publications

Journals

Rajdeep Kumar Nath, and Himanshu Thapliyal. "Smart Wristband-Based Stress Detection Framework for Older Adults With Cortisol as Stress Biomarker." IEEE Transactions on Consumer Electronics 67.1 (2021): 30-39.

Rajdeep Kumar Nath, Himanshu Thapliyal, and Allison Caban-Holt. "Machine Learning Based Stress Monitoring in Older Adults Using Wearable Sensors and Cortisol as Stress Biomarker." Journal of Signal Processing Systems (2021): 1-13.

Rajdeep Kumar Nath, Himanshu Thapliyal, Allison Caban-Holt, SP Mohanty . "Machine learning based solutions for real-time stress monitoring." IEEE Consumer Electronics Magazine 9.5 (2020): 34-41.

Himanshu Thapliyal, Rajdeep Kumar Nath, and Saraju P. Mohanty. "Smart home environment for mild cognitive impairment population: Solutions to improve care and quality of life." IEEE Consumer Electronics Magazine 7.1 (2017): 68-76.

Rajdeep Kumar Nath, Himanshu Thapliyal, " Machine Learning Based Anxiety Detection in Older Adults using Wristband Sensors and Context Feature,", 2021, Springer Nature Computer Science Journal (Accepted for Publication).

Rajdeep Kumar Nath, Himanshu Thapliyal, Travis Humble, " A Review of Machine Learning Classification Using Quantum Annealing for Real-world Applications", 2021, Springer Nature Computer Science Journal (Accepted for Publication)

Conferences

Rajdeep Kumar Nath, Himanshu Thapliyal, and Allison Caban-Holt. "Validating physiological stress detection model using cortisol as stress bio marker." 2020 IEEE International Conference on Consumer Electronics (ICCE). IEEE, 2020.

Rajdeep Kumar Nath, Himanshu Thapliyal, and Allison Caban-Holt. "Towards Photoplethysmogram Based Non-Invasive Blood Pressure Classification." 2018 IEEE International Symposium on Smart Electronic Systems (iSES)(Formerly iNiS). IEEE, 2018.

Rajdeep Kumar Nath, and Himanshu Thapliyal. "PPG Based Continuous Blood Pressure Monitoring Framework for Smart Home Environment." 2020 IEEE 6th World Forum on Internet of Things (WF-IoT). IEEE, 2020.

Rajdeep Kumar Nath, Rajnish Bajpai, and Himanshu Thapliyal. "IoT based indoor location detection system for smart home environment." 2018 IEEE International Conference on Consumer Electronics (ICCE). IEEE, 2018.

Joseph Clark, Rajdeep Kumar Nath, and Himanshu Thapliyal. "Machine Learning Based Prediction of Future Stress Events in a Driving Scenario." 2021, IEEE 7th World Forum on Internet of Things (Accepted for publication).

Rajdeep Kumar Nath, Himanshu Thapliyal, Travis S. Humble. "Quantum Annealing for Automated Feature Selection in Stress Detection.", 2021, IEEE Computer Society Annual Symposium on VLSI (Accepted for publication)

Rajdeep Kumar Nath and Himanshu Thapliyal. "Wearable Health Monitoring System for Older Adults in a Smart Home Environment.", 2021, IEEE Computer Society Annual Symposium on VLSI (Accepted for publication)