Next-Generation Healthcare with Applied Machine Learning and Quantum Machine Learning

A Dissertation Presented for the

Doctor of Philosophy

Degree

The University of Tennessee, Knoxville

Md. Saif Hassan Onim

May 2026

© by Md. Saif Hassan Onim, 2026
All Rights Reserved.

To my mother, whose sacrifices and guidance shaped who I am, and to my wife, whose patience, encouragement, and belief in me never faded–this dissertation is dedicated with love and gratitude.

Acknowledgements

I would like to express my deepest gratitude to my advisor, Dr. Himanshu Thapliyal, for his unwavering guidance, encouragement, and mentorship throughout my Ph.D. journey. Beyond his role as an advisor, his friendship, support, and trust have meant a great deal to me. His patience and high standards have shaped how I approach research and problem-solving, and I am especially grateful for everything he has taught me, from developing ideas and asking the right questions to writing, presenting, and pursuing work with integrity.

My family played one of the biggest roles in shaping who I became today. I am profoundly thankful to my mother Shahanaj Parvin for her unconditional love, sacrifices, and prayers. Her strength and support have been the foundation behind every step I have taken. I am equally grateful to my wife Jannatun Tajree for standing by me through the long days, late nights, and uncertain moments of graduate school. Her encouragement, understanding, and belief in me made it possible to stay focused and keep moving forward.

Special thanks to my lab mates Wu Yang, Joseph Clark, Tyler Cultice, Milad Nasab, JC Chin, Bhaskar Gaur, and Sounak Bhowmik for fostering a collaborative, innovative, and productive research environment.

Finally, I appreciate all the colleagues, committee members, faculties and collaborators, and friends who supported me during this process, and I thank the University of Tennessee, Knoxville, for providing an environment where I could grow both as a person and as an academic.

Abstract

Older adults often require non-intrusive monitoring tools, as their physiological and emotional responses are influenced by age-related changes, cognitive vulnerability, and psychological and social factors. Traditional approaches such as self-reports or facial recognition can be intrusive, subjective, or difficult to use for individuals with cognitive impairment. This dissertation proposes a computational framework for identifying digital biomarkers of stress, emotion, and related psychosocial states using physiological signals from wearable devices, integrating classical machine learning, deep learning, and hybrid quantum machine learning methods.

The framework begins with context-aware stress detection using both classical machine learning and deep learning approaches. Here, feature-engineered supervised learning techniques combine electrodermal activity, blood volume pulse, heart rate variability, acceleration, and skin temperature with task-level contextual information to establish a practical classification performance. Expanding from detection, time-series prediction using recurrent neural networks and long short-term memory models demonstrates the feasibility of stress forecasting and reveals temporal patterns that could support proactive interventions. Lastly, emotion-intensity estimation models based on regression techniques enable the independent modeling of emotional responses that can reduce individual variability and lay a foundation for reliable digital biomarkers.

To address the challenge of modeling complex, high-dimensional physiological relationships, the framework incorporates quantum machine learning techniques.

Physiological signals are embedded into quantum feature spaces to improve class separability and robustness to noise. Hybrid quantum models further support anomaly detection by identifying potential tampering or corruption in physiological data streams. Preliminary results indicate certain trials in these models outperform classical baselines even with modest sample sizes, showing promise for real-world healthcare and cyber-physical monitoring.

The dissertation also investigates the scalability of quantum models by examining performance as qubit counts increase and feature encodings extend to higher-dimensional representations. Systematic evaluation under Noisy Intermediate-Scale Quantum constraints analyzes trainability, generalization, and computational cost. This work establishes practical guidelines for scalable quantum-enhanced digital biomarker modeling. Together, these contributions provide a foundation for personalized, continuous physiological monitoring in older adults and at the same time, advance theoretical and empirical understanding of hybrid quantum-classical machine learning in healthcare applications.

Table of Contents

1 Introduction **1**

1.1 Current Challenges in Stress and Emotion Detection 6

1.2 Overview of Solution Strategies 7

1.3 Contributions of This dissertation 9

1.4 Outline of the dissertation 12

2 Background and Datasets **14**

2.1 Stress Detection and Prediction 14

2.1.1 Stress Protocol: Trier Social Stress Test 17

2.1.2 Ground Truth Estimation From Cortisol Concentration 18

2.1.3 Distribution of Labeled Dataset 18

2.1.4 Incorporating Context 20

2.2 Emotion Detection 20

2.2.1 Data Recording 21

2.2.2 iMotion's FEA Module 23

2.2.3 Preprocessing and Ground Truth Labeling 23

2.3 Detection and Prediction with Classical Machine Learning and CNN-Based Methods 25

2.3.1 Classical Models for Classification 27

2.3.2 Regression Methods 27

2.3.3 Time-Series and Sequential Prediction Models 28

2.4 Detection with Quantum Machine Learning 29
2.4.1 Quantum Computing 29
2.4.2 Quantum Gates 30
2.4.3 Embedding Classical Data to Quantum 31
2.4.4 Noisy Intermediate-Scale Quantum (NISQ) Computers 31
2.4.5 Quantum Benefits 33
2.4.6 Quantum-Hybrid Machine Learning Framework 33
2.4.7 Quantum Fidelity Kernels 34
2.4.8 Quantum-Hybrid Support Vector Machine 34

3 Utilizing Machine Learning for Context-Aware Digital Biomarker of Stress in Older Adults 37

3.1 Related Works and Motivation 38
3.2 Contributions 44
3.3 Data Collection, Preprocessing and Labeling 46
3.3.1 Data Collection 46
3.4 Context-Aware Stress Detection with Supervised Machine Learning . 47
3.4.1 Statistical Feature Extraction from Signal stream 47
3.4.2 Machine Learning Model 49
3.4.3 Model Training and Testing 51
3.4.4 Result Analysis 52
3.5 Context-Aware Stress Detection with CNN-Based Automatic Feature Encoder 52
3.5.1 Automatic Feature Extraction with Feature Encoder 54
3.5.2 Fully Connected Neural Network 54
3.5.3 Result Analysis 57
3.5.4 Result Comparison and Discussion 60
3.6 Predicting Stress in Older Adults with RNN and LSTM from Time Series Sensor Data and Cortisol 60

3.6.1 Proposed Time Series Prediction Model 62
3.6.2 Result Analysis and Discussion 65
3.7 Conclusion and Future Works 65

4 Emotion Detection in Older Adults Using Physiological Signals from Wearable Sensors 68
4.1 Related Works and Motivation 69
4.2 Contributions 70
4.3 Proposed Method for Emotion Detection 72
4.3.1 Machine Learning Models for Regression 72
4.3.2 Model Training and testing 76
4.4 Result Analysis and Discussion 76
4.5 Conclusion 78

5 Quantum Hybrid Support Vector Machines for Stress and Emotion Detection in Older Adults 81
5.1 Related Works and Motivation 82
5.2 Contributions 86
5.3 Proposed Method for Stress Detection 87
5.3.1 Data Collection, Preprocessing and Feature Selection 87
5.3.2 Feature Mapping with Fidelity Quantum Kernel 90
5.3.3 Training and Testing with the One-Class Support Vector Classifier 90
5.3.4 Result Analysis 91
5.4 Proposed Method for Emotion Detection 93
5.4.1 Data Preprocessing and Ground Truth Labeling 93
5.4.2 Model Training and Testing 93
5.4.3 Result Analysis and Discussion 96
5.5 Conclusion and Future Work 98

6 Detection of Physiological Data Tampering Attacks with Quantum Machine Learning **100**

6.1 Recent Works and Motivation 101

6.2 Contributions 102

6.3 Threat Model 103

6.3.1 Data Poisoning 103

6.3.2 Adversarial Perturbation 104

6.4 Proposed Tampering Detection Method 104

6.4.1 Dataset Preprocessing and Feature Extraction 104

6.4.2 Sampling and Feature Reduction 106

6.4.3 Encoding and Feature Map 107

6.4.4 Adversarial Attack with Threat Model 109

6.4.5 Tampering Detection with One-Class SVM 109

6.5 Result Analysis and Discussion 111

6.6 Conclusion and Future Work 112

7 Scalability Analysis of Quantum Models for Stress, Emotion and Data Tampering Detection **113**

7.1 Computational Infrastructure and Simulation Constraints 114

7.2 Scaling Stress Detection with QML for High Qubit Count 116

7.2.1 Result Analysis 116

7.3 Scaling Emotion Recognition with QML for High Qubit Count 117

7.3.1 Result Analysis 117

7.4 Scaling Data Tampering Detection for High Qubit Count 119

7.4.1 Result Analysis 120

7.5 Effect of Scaling For Quantum Machine Learning Framework across Physiological Domains 121

8 Conclusion and Future Research Direction **122**

8.1 Future Research Direction 123

8.1.1 Stress Modeling 123
8.1.2 Emotion Intensity Estimation 124
8.1.3 Quantum Machine Learning for Healthcare 124
8.1.4 Scalability 125

Bibliography **126**

Vita **149**

List of Tables

2.1 Emotion Baseline and Percentage Outside 1st Standard Deviation for Labeled Dataset 24

3.1 Summary of Recent Research on Older Adult Stress Detection 40
3.2 Overview of Recent Literature on Stress Prediction 42
3.3 Training Parameters for Random Forest 50
3.4 Extracted Statistical Features from Physiological Signal 50
3.5 Performance of *Random Forest* Model with and without Context . . . 53
3.6 Training Parameters for Convolutional Neural Network 56
3.7 Proposed Model Layers 56
3.8 Performance for 5-fold Cross validation 59
3.9 Confusion Matrix (weighted average of 5 folds) 59
3.10 Performance of Proposed CNN-Based Model with and without Context 59
3.11 Performance Comparison of Signal Fusion for Proposed CNN-Based Model with Machine Learning Model 59
3.12 Training Hyper-parameter and Hardware 61
3.13 Performance Metrics of Stress Prediction with LSTM and RNN for the 270th Sample (4.5 min in the future). 64
3.14 Comparison of Our Work with Contemporary Literature 66

4.1 Overview of Recent Literature on Emotion Detection from Physiological Signals 71

4.2 Hyper-parameters Used to Train Machine Learning Models 77
4.3 r^2 Scores for Various ML Models (Higher is better) 79
4.4 MSE for Various ML Models (Lower is better) 79

5.1 Recent Literature on Emotion Detection 85
5.2 Performance Metrics of Quantum SVM Model and Classical SVM Based on Feature (and qubit) counts for Stress Detection. 92
5.3 Hyper-parameters Used to Train Machine Learning Models 95
5.4 Performance of Classical ML Models for 2000 Samples (avg of 5 folds) 97
5.5 Performance of Hybrid Quantum SVM Model with Belis Kernel (avg of 5 folds) 97

6.1 Description of Dataset 105
6.2 Average Performance of Dot Product Kernel SVM for 5-fold Data Before and After Adversarial Attack 110
6.3 Detection Accuracy for Test Set of Proposed Framework after 5-fold Cross validation 110

7.1 HPC Resources and Simulation Details for Quantum Model Scalability Experiments 115
7.2 Performance Comparison of Stress Detection for High Qubit Configurations 115
7.3 Comparison of Emotion Recognition Performance: 4-Qubit vs. 8-Qubit Configurations 118
7.4 Impact of Qubit Scaling on Data Tampering Detection Accuracy . . . 118

List of Figures

1.1 Acute and Chronic Stress [1] 2
1.2 Human Response to Stress Curve [2] 2
1.3 Mean Emotion-wise Classification Accuracy Across Age Groups and Developmental Trajectories for the Emotion [3] 5
1.4 How Participants Perceived Neutral or Happy/Angry Facial Expressions of Young and Older Adults [4] 5

2.1 Cognitive Performance vs AD Continuum due to chronic stress [5] . . 16
2.2 3 Labels for Cortisol Concentration of a Participant 16
2.3 2 Labels for Cortisol Concentration of a Participant 16
2.4 TSST Protocol; PS: Pre-Stress, AS: Anticipatory Stress, SA: Speech Assignment, MA: Mental Arithmetic 19
2.5 Distribution of Projected 2D labeled dataset with FLD 19
2.6 Contextual Information for Stress Detection in Recent Literature . . . 19
2.7 Facial Muscles Movements Captured During FEA 22
2.8 Arousal-Valance Axis 22
2.9 2D Projection of the Labeled Dataset with PCA Showing the Cluster Centers where Each Color Represents Different Emotion 24
2.10 A Generic Classical Machine Learning Pipeline. (WTA: Winner-Take-All, SVM: Support Vector Machines, RF: Random Forest, XGB: X-Gradient Boosting, DT: Decision Tree, MPP: Maximum Posterior Probability, KNN: K-Nearest Neighbor, MLP: Multi-Layered Perceptron) 26

2.11 Quantum Gate Symbols 32
2.12 Bloch Sphere Showing a State $|\psi\rangle$ with Parameterized θ and φ 32
2.13 Quantum Hybrid Machine Learning Framework [6] 35
2.14 Fidelity Quantum Kernel $k(x_i, x_j) = |\langle 0|U^\dagger(x_i)U(x_j)|0\rangle|^2$. (The vectors x_i and x_j are N-dimensional features encoded into quantum states using S_{x_i}, with optional entangling gates $G(x_i)$. The circuit applies $U(x_i)$, followed by its inverse $U^\dagger(x_j)$, and measures the overlap to compute the kernel) 35

3.1 Empatica Wristband for Data Collection 42
3.2 Workflow of Stress Detection with Machine Learning Model(Random Forest) 48
3.3 Workflow of Stress Detection with CNN-Based Model 55
3.4 Workflow of Proposed Stress Prediction Model 61
3.5 Predicted Stress Label Sample vs Accuracy 64

4.1 Proposed Emotion Detection Method 73
4.2 Correlation Heatmap of the Features and Emotion of Dataset 75

5.1 Pipeline of the Hybrid-Quantum SVM Anomaly Detection 88
5.2 Quantum Circuit for Feature Mapping with 8 qubits 88
5.3 2D Projection of Test Kernel Matrix with Classical and Quantum Kernel. (*blue* represents Baseline or Normal class and *orange* represents Anomaly or Stress class) 92
5.4 Proposed Emotion Detection Method with QML 94

6.1 Workflow of Proposed Hybrid QML model for Tampering Detection . 105
6.2 Kernel Computation Circuit. $(0-5)$ are the qubits, H is the Hadamard gate, R_x and R_y are the Rotation gates, $\oplus$ denotes the C-NOT gates and † represents the Reverse operations 108

6.3 2D Projection of EPHNOGRAM Dataset Before and After Data Poisoning Attack. (*blue* and *orange* represent two different classes . . 108

Chapter 1

Introduction

Stress and emotion play crucial roles in determining physical, cognitive, and psychological health, especially for older adults who experience heightened vulnerability to physiological and cognitive changes. Stress is shaped by social, economic, psychological, and physical factors, and is widely regarded as an indicator of overall quality of life [7]. A growing body of research demonstrates that chronic stress increases susceptibility to a range of illnesses, including cardiovascular disease, immune dysfunction, metabolic disorders, and depression [8, 9, 10]. In older adults, prolonged stress exposure is also tied to increased cognitive decline and may contribute to the progression of Alzheimer's disease and related dementias (ADRD) [11].

While moderate stress can motivate adaptation, growth, and creativity, excessive or chronic stress has well-documented negative consequences. Prolonged high stress levels contribute to serious health problems, reduce work capacity, and degrade quality of life [7, 12, 13]. Large-scale population studies highlight the magnitude of this issue. Recent global data covering 1.53 million individuals across 113 countries from 2009 to 2021 reveal about a 25% to 31% rise in self-reported emotional distress, such as worry, sadness, stress, or anger, at the population level [14]. In 2012, 87% of adults in the United States and Great Britain identified stress as a major health concern. Altogether 63% reported knowing someone who experienced stress-related health



Figure 1.1: Acute and Chronic Stress [1]

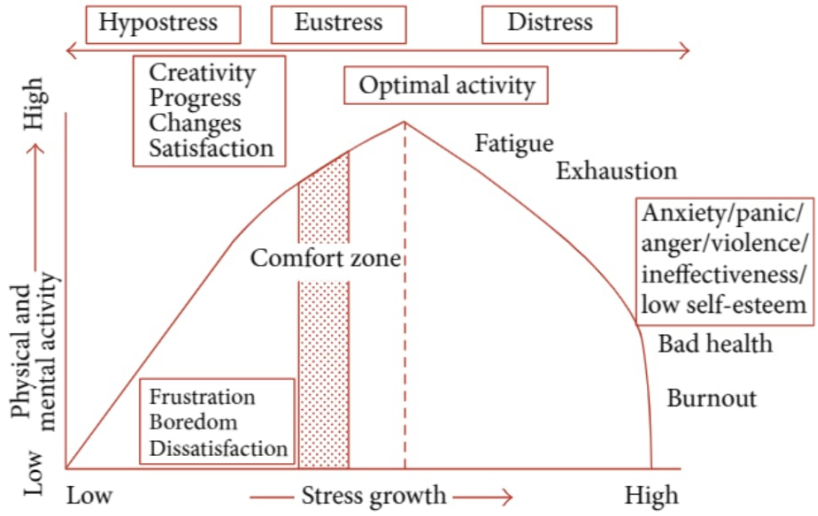


Figure 1.2: Human Response to Stress Curve [2]

effects and more than 40% reported experiencing stress themselves [13]. A Canadian survey similarly found that more than 60% of highly stressed employees attributed their stress primarily to work-related factors [15]. Socioeconomic disparities including differences in income, education, and occupation have been shown to account for susceptibility to mood and anxiety disorders [7, 16]. Some common stressful activities are shown in Figure 1.1 separated based on acute and chronic nature. When generalized human reaction to stress in terms of physical and mental activity follows a bell-shaped curve shown in Figure 1.2 provided by a study back in 1975 and reported in another recent article [2]. It shows how increasing stress initially enhances performance up to an optimal point before leading to fatigue, distress, and eventual burnout as stress levels become excessive.

Certain occupational groups experience disproportionately high levels of stress. Healthcare and education professionals, civil service workers, and maritime personnel consistently report elevated stress burdens [17, 18]. Among seafarers, for example, an estimated 28.3% to 63.5% experience significant work-related stress [19]. Contributing factors include chronic sleep deprivation, poor sleep quality, long and irregular working hours, high workload, circadian rhythm disruptions, noise, heat, vibration, limited mobility, dehydration, climatic and time-zone changes, lack of access to medical professionals, social isolation, and disruptions to personal and family life.

Similarly, emotion also shapes well-being, cognitive functioning, and day-to-day behavior. Its significance becomes even more prominent in older adult populations. Emotional states influence attention, memory, decision-making, and physiological regulation, and disruptions in emotional processing are strongly linked to increased risk of depression, anxiety, and accelerated cognitive decline [20]. In aging populations, emotional changes often coincide with shifts in social roles, reduced mobility, greater medical burden, and increased exposure to loneliness or social isolation, all of which heighten vulnerability to negative emotional experiences [21]. Epidemiological evidence shows that mood and anxiety disorders remain highly

prevalent among older adults, with emotional irregularity serving as both a predictor and consequence of age-related health challenges [22].

These issues are further reinforced in individuals living with or at risk for Alzheimer's disease and related dementias (ADRD), where emotional symptoms such as agitation, apathy, and heightened stress response commonly emerge with cognitive impairment. Figure 1.4 shows a study by Hass *et al.* [4] claiming that emotional expressions of old faces are perceived as more positive and than in young adults. Physiologically, emotional states in older adults are reflected in measurable autonomic changes, including variations in heart rate variability, electrodermal responses, and peripheral blood flow, all of which provide important indicators of emotional arousal [21]. However, these may differ substantially from younger populations due to age-related changes in neural circuitry, autonomic balance, and stress-resilience mechanisms. Some studies show that they are even harder to detect compared to the younger adults as well, as shown in Figure 1.3. These complexities underscore the need for objective, continuous, and unobtrusive emotion monitoring methods that can operate in naturalistic settings and support early identification of emotional or cognitive decline in older adults.

Given the substantial impact that both stress and emotional well-being have on the health trajectories of older adults, there is a growing need for assessment approaches that can capture these states in everyday life. Monitoring with conventional techniques is difficult given the complex and dynamic effects of stress and emotions brought on by age, individual differences, and environmental influences. These challenges highlight the limitations of traditional assessment techniques and underscore the need for objective, scalable, and unobtrusive methods capable of detecting stress and emotional responses outside controlled laboratory environments. Traditional stress and emotion assessment methods rely on self-report questionnaires, structured behavioral tasks, or camera-based facial expression analysis [23, 24]. While effective in controlled environments, these approaches face limitations in real-world settings. They require active user participation, may be biased by memory or

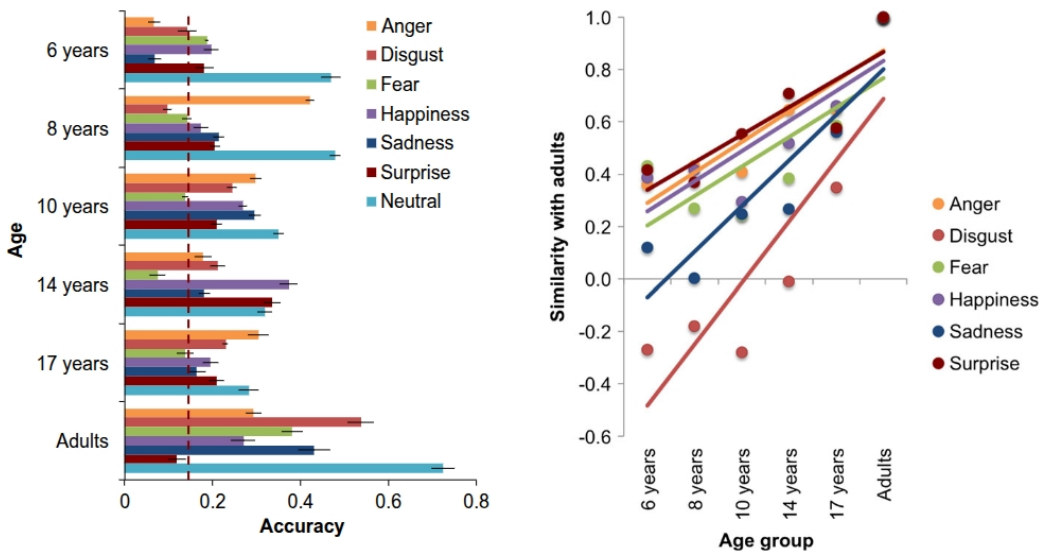


Figure 1.3: Mean Emotion-wise Classification Accuracy Across Age Groups and Developmental Trajectories for the Emotion [3]

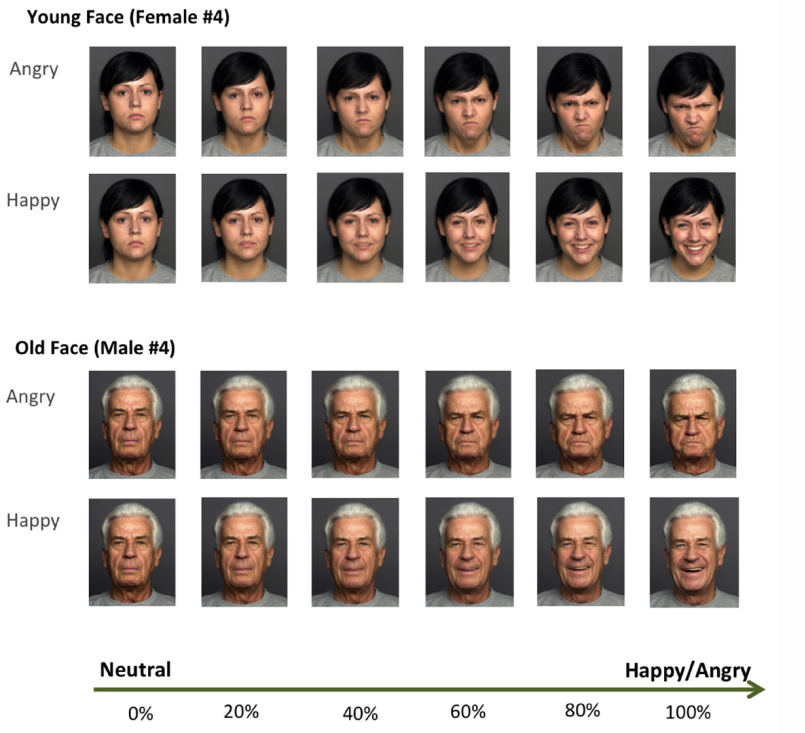


Figure 1.4: How Participants Perceived Neutral or Happy/Angry Facial Expressions of Young and Older Adults [4]

subjective reporting, raise privacy concerns, and can be difficult for older adults or individuals with cognitive impairments to use consistently.

Recent advancements in wearable physiological sensors such as the Empatica E4 and Shimmer3 have enabled researchers to collect signals including electrodermal activity (EDA), blood volume pulse (BVP), accelerometer data, skin temperature, and heart rate variability (HRV) in real time [25]. These signals correlate strongly with autonomic nervous system activity and thus provide unobtrusive, continuous indicators of stress and emotional arousal.

1.1 Current Challenges in Stress and Emotion Detection

Early work in stress detection primarily relied on controlled laboratory protocols such as the Trier Social Stress Test (TSST), where ground-truth labels were assigned manually based on predefined task segments [23]. Classical machine learning approaches in these settings typically used statistical features such as means, variances, peaks, and spectral descriptors and applied classifiers including support vector machines (SVMs), k-nearest neighbors (k-NN), random forests, and logistic regression [25]. While effective in structured laboratory environments, these methods often struggle to generalize to real-world conditions and diverse populations.

Emotion detection followed a similar trajectory. Early methods primarily relied on facial expression analysis, speech processing, and self-report instruments [23, 24]. These techniques, although informative, raise privacy concerns, require voluntary participation, and often perform poorly for older adults or individuals with cognitive impairments. With the emergence of wearable sensors, researchers increasingly turned to physiological-signal-based emotion detection using modalities such as electrodermal activity (EDA), heart rate variability (HRV), blood volume pulse (BVP), and skin temperature [26, 27]. These sensing modalities enable continuous, non-intrusive

monitoring, but they also introduce more modeling challenges. Despite progress, current stress and emotion detection methods face several critical issues:

- **High variability:** Physiological responses differ widely among individuals, particularly in older adults, due to age-related autonomic changes, comorbidities, and medication effects.
- **Noisy and temporally complex signals:** Wearable sensor data are prone to motion artifacts, environmental noise, and nonlinear temporal patterns that classical models cannot fully capture.
- **Dataset imbalance and limited scalability:** Real-world datasets often have fewer positive stress or high-intensity emotional events, making training difficult.
- **Computational constraints:** Deep learning models capable of capturing nonlinear patterns are computationally expensive and challenging to deploy on resource-limited, real-time platforms typical of older-adult monitoring systems.

These challenges motivate the exploration of more advanced computational paradigms capable of modeling noisy, multimodal, and nonlinear physiological signals while operating efficiently in practical, real-world environments.

1.2 Overview of Solution Strategies

This section provides a high-level overview of the core methodological strategies that guide the proposed research. These strategies address the challenges outlined in Section 1.1, particularly the difficulty of extracting reliable stress and emotion biomarkers from noisy, multimodal physiological signals in older adults and the computational requirements of modeling high-dimensional temporal data. The solution strategies are as follows:

Multimodal Digital Biomarker Development

Multimodal digital biomarkers integrate multiple physiological signals to capture stress and emotion related changes more thoroughly than single sensor stream. Wearable signals such as electrodermal activity (EDA), blood volume pulse (BVP), heart rate variability (HRV), skin temperature (ST), and respiration carry related information about affective states. Combining these signals allows the model to reduce ambiguity inherent in individual modalities and capture more detailed physiological signatures. In this work, we:

- proposed continuous, passive, and non-intrusive monitoring suitable for older adults and individuals with cognitive vulnerabilities.
- engineered multimodal features to derive reliable stress and emotion biomarkers that are physiologically grounded and appropriate for real-world deployment.

Temporal Deep Learning Models

Physiological responses unfold over time, which makes time series modeling essential for accurate inference and prediction. We, therefore, investigated here:

- if recurrent neural networks (RNNs) and long short-term memory (LSTM) architectures can model long-range temporal dependencies within physiological patterns
- how they capture transitions between stress states and predict future stress levels with variable-length time windows.

Hybrid Quantum-Classical Learning

Quantum computing introduces a fundamentally different approach for representing and learning from complex data distributions. In this work, we investigated:

- whether hybrid quantum learning can reduce model complexity by offloading high-dimensional feature transformations onto quantum circuits.
- how they improve separability of overlapping physiological patterns through quantum feature spaces.
- opportunities for scalable, next-generation computational tools for healthcare applications.

1.3 Contributions of This dissertation

This dissertation presents several technical, methodological, and application-driven contributions across stress detection, emotion intensity estimation, and physiological data integrity using classical machine learning, deep learning, and quantum machine learning. They include dataset processing, feature engineering, model design, contextual reasoning, and quantum-enhanced classification for complex physiological data collected from wearable sensors. The primary contributions of this work are listed below:

- **Context-aware stress detection using multimodal physiological signals:** This dissertation introduces a systematic and comprehensive classical machine learning framework to detect stress using physiological features extracted from electrodermal activity (EDA), blood volume pulse (BVP), skin temperature (ST), and accelerometer measurements collected during the Trier Social Stress Test (TSST). A key novelty of this contribution is the integration of *contextual task information* from the TSST such as baseline, anticipation, speech, and arithmetic phases, into the classification process. By treating contextual information as an additional modality fused with physiological features, the proposed model demonstrates large performance gains in accuracy and F1 score compared to non-contextual baselines. This approach shows

improvements across five-fold cross-validation and highlights the importance of context and significantly reduces false positives in stress classification.

- **CNN-based automatic feature encoding for stress classification:** To address the limitations of manually engineered features and traditional models, this dissertation develops a deep learning-based automatic feature encoder using convolutional neural networks (CNNs). The model operates directly on raw or minimally processed physiological time-series signals and learns temporal and spatial features without the need for handcrafted statistical computations. The CNN architecture incorporates multiple convolutional layers to detect local patterns, max-pooling layers for downsampling, and fully connected layers for feature abstraction. This enables the model to capture nonlinear interactions among physiological modalities. Experiments conducted across multiple TSST segments demonstrate that this CNN-based encoder outperforms classical methods, particularly in conditions with overlapping stress-label boundaries.

- **Stress prediction with time-series models:** This dissertation also contributes a modeling framework for predicting stress using sequential physiological data collected from wearable sensors. Unlike static feature-based classifiers that rely on aggregated statistics over fixed windows, time-series models capture the dynamics of physiological responses. They can be changes in heart rate variability, electrodermal fluctuations, and nonlinear trends in blood volume pulse, and so on. We develop and evaluate recurrent neural architectures, including LSTM-based models, to assess their ability to forecast upcoming stress states rather than simply classify current instances. These models reveal how temporal dependencies, short-term physiological transitions, and pre-stress patterns can be leveraged to anticipate stress events before they are fully expressed. This contribution highlights the potential of predictive physiological modeling for early stress intervention in older adults.

- **Emotion intensity estimation from wearable physiological data using classical regression models:** This dissertation advances a framework for estimating continuous emotion intensity by leveraging multimodal physiological features. A range of regression models, including Ridge Regression, Random Forest Regression, Gradient Boosting Regression, and multilayer perceptron (MLP) regression, are evaluated to determine how classical learning methods perform on real-world datasets with complex physiological responses. This contribution demonstrates that physiological-signal-based emotion modeling can be accomplished without intrusive audio or video modalities. This supports privacy-preserving affective computing applications for home monitoring, cognitive health assessment, and behavioral research in older adult populations.

- **Quantum hybrid SVM for stress and emotion detection using quantum kernel methods:** To address the challenges posed by high-dimensional and nonlinear physiological datasets, this dissertation extends the classical algorithms to quantum-enhanced models trained using quantum kernel functions generated by parameterized quantum circuits. These circuits serve as expressive feature maps capable of embedding the input data into high-dimensional Hilbert spaces and enable the computation of similarities between data points using quantum superposition and entanglement. The contribution includes the design of multiple quantum feature map architectures, evaluation of kernel matrices on simulated noisy intermediate-scale quantum (NISQ) devices, and benchmarking against classical SVM models. This contribution demonstrates one of the earliest applications of quantum-enhanced learning for physiological signal analysis and provides practical insights into designing circuits that balance hardware constraints.

- **Physiological data tampering detection using quantum machine learning for healthcare IoT systems:** Beyond stress and emotion modeling, this dissertation applies quantum kernel methods to detect data tampering

and adversarial manipulation in physiological sensor streams. Thus, our work addresses an increasingly important cybersecurity concern in wearable and IoT-based healthcare systems. We develop a quantum-kernel anomaly detection framework that identifies irregularities and unnatural patterns introduced through data alteration attacks. By constructing quantum feature maps capable of amplifying differences between authentic and adversarially altered data, the proposed system achieves strong precision-recall tradeoffs in detecting data tampering. The contribution includes detailed comparisons between classical and quantum anomaly detectors, analysis of quantum circuit depth and entanglement on detection accuracy, and evaluation under various simulated attack conditions. This work demonstrates that quantum models can play a dual role in healthcare IoT that ensures data integrity in healthcare monitoring.

- **Scalability of quantum models in healthcare:** We analyzed performance trends as the number of qubits increases and evaluated different angle encoding strategies, including one-feature-per-qubit and two-features-per-qubit mappings for a higher count of qubits. The study examines how qubit scaling influences performance metrics and training stability. Our experiments reveal that performance gains are not strictly linear with increased qubit counts. These results offer actionable design guidance for implementing scalable quantum models in digital biomarkers focused on healthcare.

1.4 Outline of the dissertation

The dissertation is organized as follows. Chapter 1 introduces the stress and emotion detection problem, outlines the main contributions, and presents the overall structure of the work. Chapter 2 provides background, describes data collection protocols, and reviews relevant literature. Chapter 3 presents context-aware stress detection and prediction using classical and CNN-based machine learning models. Chapter 4

examines emotion intensity estimation from physiological signals, evaluates multiple regression models, and discusses implications for real-world monitoring in older adults. Chapter 5 introduces quantum hybrid support vector machine models for stress and emotion detection, detailing circuit design, quantum feature maps, and comparisons with classical approaches. Chapter 6 extends quantum methods to physiological data tampering detection through a quantum-kernel anomaly detection framework for healthcare IoT systems. Chapter 7 presents the scalability analysis of quantum models for stress, emotion, and data tampering detection as the final part of the dissertation. Lastly, Chapter 8 concluded the dissertation with possible future research direction.

Chapter 2

Background and Datasets

In this section, we will provide a background to stress and emotion detection, relevant protocol, data collection, and quantum computing. We will also cover the basics of classical computing and the effects of quantum computing, both as an implementation tool and a defense against adversaries.

2.1 Stress Detection and Prediction

Stress monitoring plays a crucial role in healthcare, as it impacts individuals' cognitive functions and decision-making abilities [28]. It significantly influences work efficiency, and prolonged exposure to excessive stress and negative emotions can detrimentally affect both physical and mental well-being [29]. This is especially pertinent for individuals in specific professions and specific age groups. During the pandemic alone, approximately 40% of adults in the United States have reported experiencing symptoms of stress, anxiety, or depressive disorder, and older adults are more susceptible to the impacts of stress compared to other age groups [30]. This has resulted in massive unemployment and increased behavioral health disorders. Due to a lack of connectivity, older adults were cut off from their families, friends, essential services, and communities. This resulted in about 4 in 10 adults in the U.S. having reported symptoms of stress, anxiety, or depressive disorder. Stress can threaten

a person's physical and emotional well-being. It can lead to increased risks of emotional and physical diseases, including immune system problems and damage to the cardiovascular system and metabolic system, *etc*. [31]. Further, stress is linked to cognitive impairments, which over time may cause the onset of Alzheimer's disease in older persons, along with physiological irregularities [11]. Figure 2.1 illustrates that the Alzheimer's Disease (AD) spectrum, encompassing a gradual progression from subtle cognitive decline through various stages to dementia, may be influenced by chronic elevations in cortisol concentration [5]. Stress-induced physiological responses, which are measurable as neurophysiological biomarkers, can offer valuable insights for understanding and potentially predicting AD progression. So, it's crucial to periodically assess and control one's stress levels to lessen the negative impacts of stress. Making an accurate diagnosis based on historical records can also be aided by frequently monitoring and reporting stress profiles. This can encourage independence, ease the burden, and ultimately enhance the quality of life.

Digital biomarkers for stress refer to measurable and quantifiable indicators of stress levels that are collected through various sensors. These biomarkers provide insights into an individual's physiological or behavioral responses to stress, offering a more objective and real-time assessment compared to traditional self-reporting methods. Commonly employed digital biomarkers for detecting stress include Electrodermal Activity (EDA), Photoplethysmography (PPG), respiration, Electroencephalogram (EEG), and Skin Temperature (ST) [29]. Multiple studies are showing the differences in digital biomarker responses in different moods or stress levels [32, 33, 34].

Stress states can also manifest in behaviors such as eye movements, facial expressions, sleeping, and speech patterns [35]. These measures have been extensively utilized in laboratory settings for stress detection. Physiological signals are more practical to collect and offer greater objectivity than behavioral indicators. The advent of wearable sensors has provided a solution, enabling the use of wrist-worn devices for portable, inconspicuous, and non-invasive collection of physiological

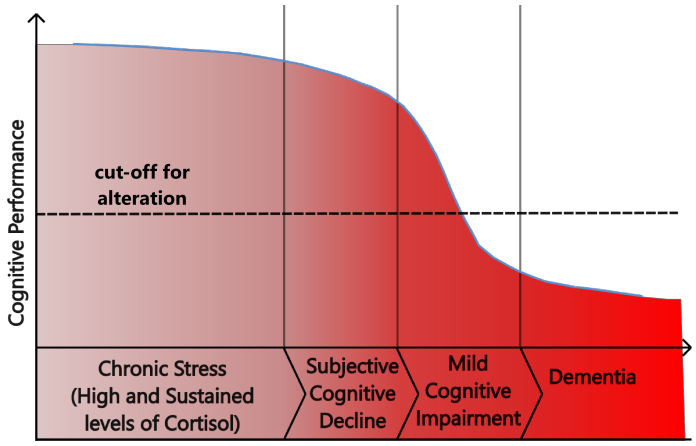


Figure 2.1: Cognitive Performance vs AD Continuum due to chronic stress [5]

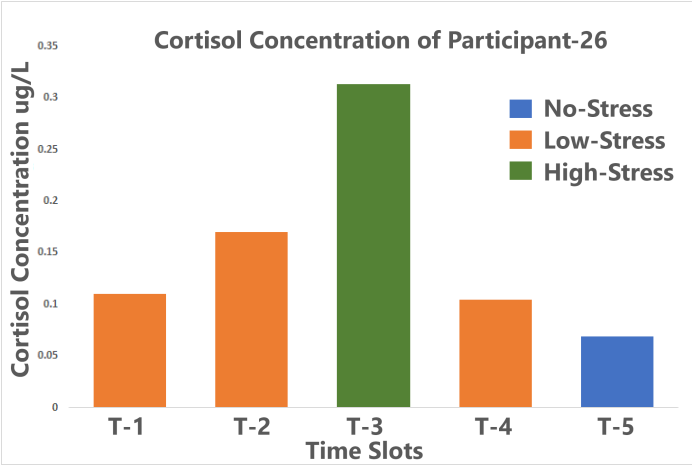


Figure 2.2: 3 Labels for Cortisol Concentration of a Participant

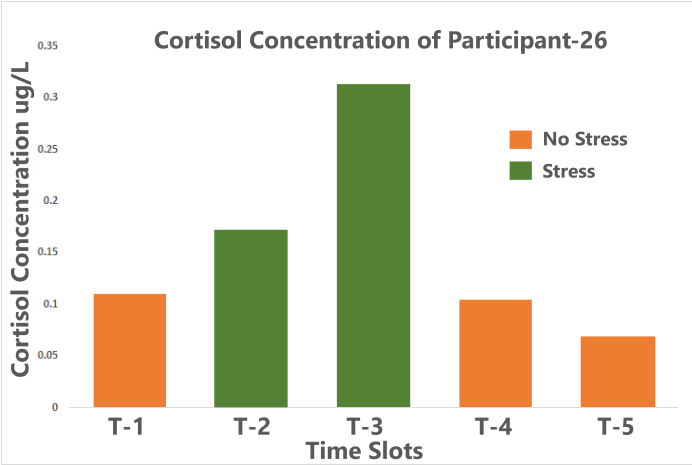


Figure 2.3: 2 Labels for Cortisol Concentration of a Participant

signals. Exercise, walking, and running, on the other hand, can generate a stress-like reaction even though they are not detrimental to health. To address the problem, context-based machine learning algorithms that combine context with digital biomarkers are being developed to distinguish between different sources of stress [36].

2.1.1 Stress Protocol: Trier Social Stress Test

The Trier Social Stress Test (TSST) protocol is a widely recognized and established experimental framework and is distinguished for its capacity to stimulate stress within a context that emulates naturalistic conditions [37]. The procedural sequence of the TSST encompasses distinct phases, as shown in Figure 2.4. It consists of the waiting interval, pre-stress phase, stress induction period, and subsequent recovery phase. As part of the preparatory stage, participants engage in the completion of demographic questionnaires. After the waiting period, the pre-stress (PS) phase starts, during which the base measurements are acquired. The collective duration of the waiting interval and the pre-stress phase comprises a duration of 20 minutes (T1-T2). Then the pre-stress and stress phases span a duration of 20 minutes (T2-T3). The stress-inducing session is partitioned into three discrete segments: an initial 10-minute span designated as the anticipatory stress phase (AS), succeeded by a 5-minute interval encompassing speech and cognitive arithmetic exercises (M). During the anticipatory stress phase (AS), participants are tasked with a continuous discourse on a designated topic for 5 minutes while being observed. Then they are engaged in the cognitive arithmetic exercise that requires resolving basic addition and subtraction equations. Importantly, the complexity of the cognitive tasks increases progressively with each accurate solution. The study ends with the inclusion of two sequential 20-minute recovery intervals (T3-T5). Throughout the course of the experiment, salivary samples were collected on five occasions corresponding to the time points denoted as T1, T2, T3, T4, and T5.

2.1.2 Ground Truth Estimation From Cortisol Concentration

The participants' saliva samples were dispatched to the laboratory for cortisol concentration analysis. We classified cortisol levels into two distinct categories, namely *No-Stress* and *Stress*. And for three labels, we named those *No-Stress*, *Low-Stress,* and *High-Stress*. They served as the reference for addressing the multiclass classification problem. For two-class classification, we mapped the highest 2 values as *Stress* and the rest as *No-Stress,* as shown in Figure 2.3. For three-class classification, we designated the minimum value among the samples taken at time points T1, T2, and T5 as *No-Stress*, with the other two being categorized as *Low-Stress*. The higher of the two samples collected at time points T3 and T4 was labeled as *High-Stress*, while the remaining one was classified as *Low-Stress*. Ultimately, these labels were expressed as numeric integers [38]. Figure 2.2 shows an example of cortisol concentration for a participant from the dataset. According to the algorithm mentioned before, the assigned labels would be *No-Stress*: T-5, *Low-Stress*: T-1, T-2, and T-4, and *High-Stress*: T-3.

2.1.3 Distribution of Labeled Dataset

The labeled data was then projected with Fisher's Linear Discriminant (FLD) in Figure 2.5. It is a dimensionality reduction method that uses a projection vector $\vec{w_p}$ to reduce the dimension of the dataset from f to $n-1$ shown in Eq. (2.1).

$$\underset{M\times(n-1)}{X_{FLD}} = \underset{M\times f}{X^T_{original}} \cdot \underset{f\times(n-1)}{\vec{w_p}} \tag{2.1}$$

Here M is the length of data, f is the original number of features, and n is the number of classes. It shows that our new ground truth labeling successfully differentiates the *No-Stress* and *High-Stress* classes with some overlap between *Low-Stress* and *High-Stress*. There are also minor data points that can be marked as

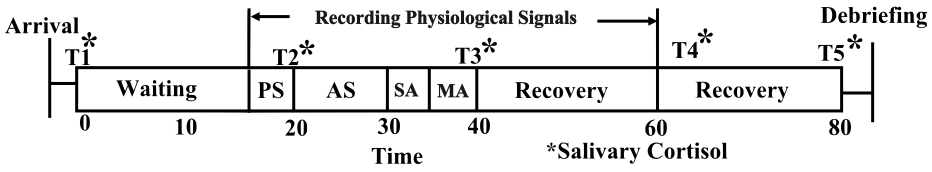


Figure 2.4: TSST Protocol; PS: Pre-Stress, AS: Anticipatory Stress, SA: Speech Assignment, MA: Mental Arithmetic

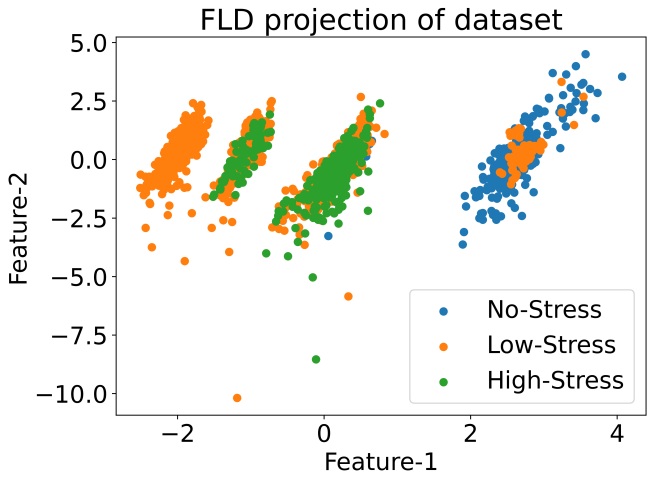


Figure 2.5: Distribution of Projected 2D labeled dataset with FLD

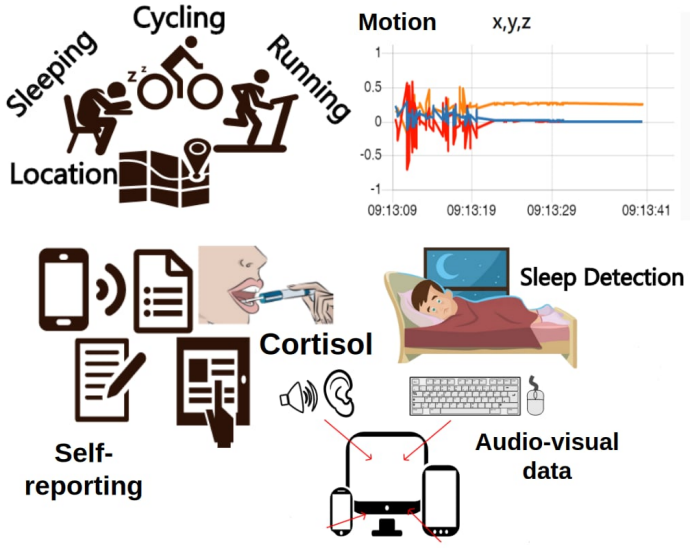


Figure 2.6: Contextual Information for Stress Detection in Recent Literature

outliers. These data points will be misclassified with any linear classifier unless some external information or context is provided to deal with them.

2.1.4 Incorporating Context

Context typically refers to the information surrounding a particular event or situation that can be leveraged to make better predictions or decisions. This surrounding information can include a wide range of factors depending on the application. In this research, our contextual information is derived from the stress levels of participants before, during, and after the TSST. It's important to acknowledge that some participants may have arrived at the lab already experiencing stress levels higher than what was induced during the TSST, while others may exhibit varying rates of stress level change following the TSST, leading to several challenges for our classifier. Primarily, these variations can result in an increased false positive rate, as the prior probability after ground truth estimation is not uniform across all classes, thus causing dataset imbalance. Additionally, the duration of the relaxed period may vary among individuals. To mitigate these unexpected outcomes and enhance decision-making, we incorporate supplementary contextual information directly into our dataset. These context features provide numeric values analogous to the ground truth labels, and their assignment is based on the laboratory conditions under which the data were collected. Our work will highlight the performance of the proposed framework in both cases: with and without context. Some of the contextual information incorporated in different literature is shown in Figure 2.6.

2.2 Emotion Detection

Understanding emotion is vital to understanding human behavior, especially in older adults, as emotional well-being is significantly related to cognitive and physical health. Research has shown that changes in emotional patterns can serve as early indicators

of neurodegenerative diseases such as Alzheimer's disease and related dementia (ADRD) [39]. They also help diagnose mental health conditions such as PTSD, depression, and anxiety disorders in veterans and other vulnerable populations. Traditional emotion recognition approaches are primarily based on facial expression analysis, speech recognition, and behavioral assessments [40, 41]. They often require camera-based monitoring or active user participation. Although effective, these methods raise concerns about privacy, intrusiveness, and real-world applicability.

Recent advancements in wearable technology and physiological sensing have introduced a promising alternative to detect mental health conditions and relevant cognitive impairments [36, 42, 43]. Researchers employ both obtrusive and non-obtrusive instruments to gather physiological signals that serve as biometric indicators [44, 45, 46]. Beyond emotion recognition, previous studies have explored stress detection using these physiological signals. For example, research that uses EDA and HRV has shown promising results in identifying acute stress responses in real-time settings. Other studies have used Empatica E4 to assess the cognitive load, anxiety, and stress during educational tasks [38, 47, 48]. These works reveal strong correlations between physiological changes and fluctuations in mental state and highlight the versatility of physiological sensors.

2.2.1 Data Recording

For recording the physiological sensor data, two commercially available smart devices called the Empatica E4 and Shimmer3 GSR+ were used. These devices have EDA, PPG, and ST sensors built in. The recording mode keeps collecting data that can be synchronized after each use. Features collected from the sensors are yaw, pitch, and roll from the accelerometer, temperature, internal ADC voltage, GSR resistance, heart rate, and GSR conductance.

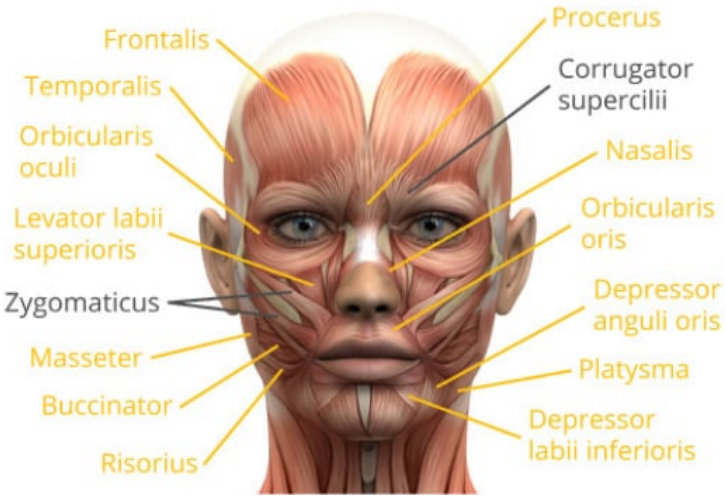


Figure 2.7: Facial Muscles Movements Captured During FEA

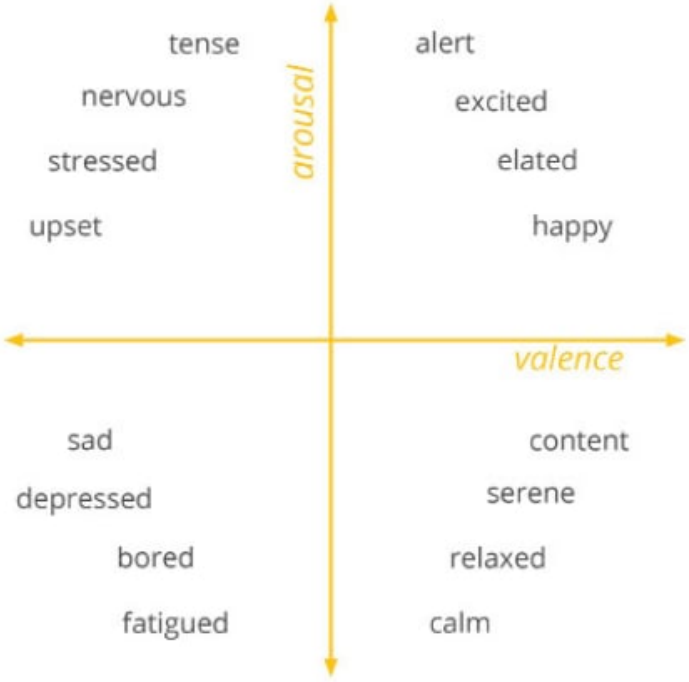


Figure 2.8: Arousal-Valance Axis

2.2.2 iMotion's FEA Module

The iMotions FEA (Facial Expression Analysis) module leverages deep learning (DL) algorithms to automatically detect and interpret facial expressions from video input, providing robust, real-time analysis of emotional states. Trained on large datasets of human faces, these DL models can accurately identify subtle facial movements and map them to basic emotions such as joy, anger, sadness, fear, surprise, disgust, and contempt. In addition to categorical emotion labels, the module also outputs emotion intensity scores, offering a quantitative measure of how strongly each emotion is expressed. The slight muscle movements that contribute to any specific emotions are recorded and taken into account while measuring the intensity as shown in Figure 2.7 and placed in the arousal-valence axis shown in Figure 2.8. This intensity metric adds valuable granularity, enabling researchers to distinguish between weak and strong emotional responses. When combined with other physiological data streams in the iMotions platform, such as eye tracking or electrodermal activity (EDA), the FEA module provides a comprehensive and synchronized view of affective and cognitive states in both controlled and naturalistic environments.

2.2.3 Preprocessing and Ground Truth Labeling

The dataset is preprocessed for labeling after data collection. A 60-sample moving average window was used to scan the time series data during preprocessing. This assisted in eliminating unwanted artifacts and noise. To prevent data leakage, the train and test sets of participants are later divided and normalized independently using z-score normalization. In our study, ground truth labeling for emotion recognition is established using iMotion's Facial Expression Analysis (FEA) module [49], which employs an advanced implementation of the Facial Action Coding System (FACS). This is a widely used framework in affective computing that can identify facial muscle movements known as Action Units (AUs). Each combination of such action units can correspond to specific emotional expressions. The FEA module

Table 2.1: Emotion Baseline and Percentage Outside 1st Standard Deviation for Labeled Dataset

Emotion	**Baseline**	**% outside 1st std**	**Emotion**	**Baseline**	**% outside 1st std**
Joy	-0.6629763	24.28	Sadness	-0.515366	27.38
Anger	0.3638698	41.11	**Neutral**	-0.0408471	34.96
Surprise	-1.859378	7.61	**Positive**	-0.6629763	24.28
Fear	-0.8419589	15.12	**Negative**	0.3638698	18.49
Contempt	-0.1459146	33.83	Confusion	0.809446	46.93
Disgust	-0.3184329	21.58	Frustration	0.5876518	42.75

⋆ In This work, we will detect only *Neutral, Positive* and *Negative* emotion

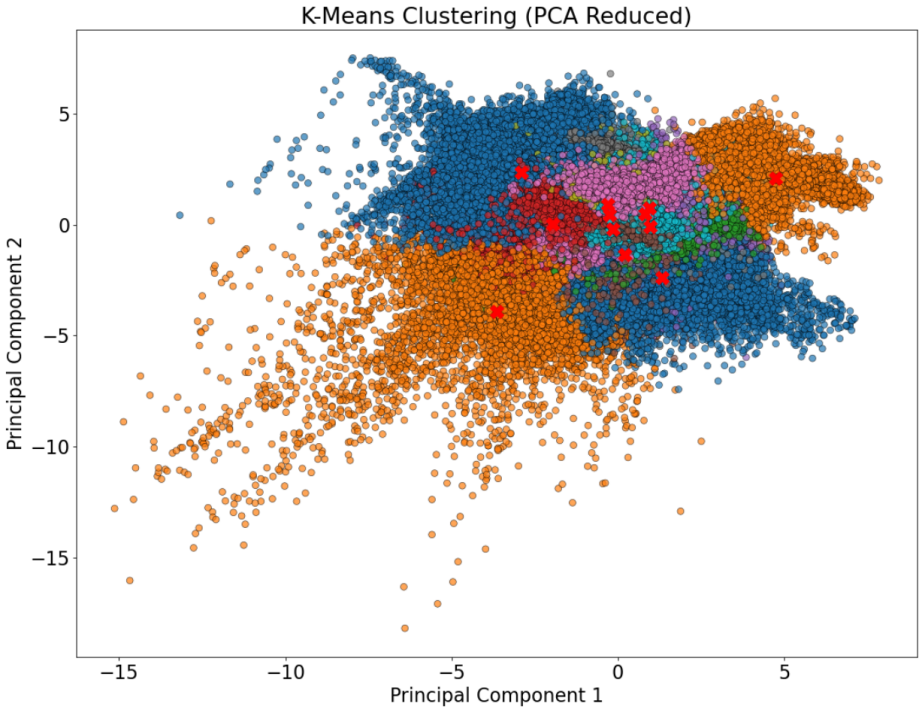


Figure 2.9: 2D Projection of the Labeled Dataset with PCA Showing the Cluster Centers where Each Color Represents Different Emotion

leverages computer vision and deep learning algorithms to analyze these AUs in real time.

This automated analysis ensures a standardized, objective method for labeling emotions, reducing potential bias associated with self-reporting. By synchronizing physiological sensor data (EDA, BVP, and IBI) with facial expression-derived emotion labels, we created a robust dataset that enables training and validation of our edge-based emotion recognition framework. After the labeling, a distribution of emotion intensity is shown in Table 2.1. Again, Figure 2.9 shows a 2D projection of all the emotion classes after labeling. The centroid of the cluster shows how all the emotional states can overlap and are hard to distinguish. In this work, we plan to detect only 3 sets of emotions: *Positive, Negative* and *Neutral*.

2.3 Detection and Prediction with Classical Machine Learning and CNN-Based Methods

Classical machine learning refers to a broad set of algorithms that use feature-based data representations to learn discriminative or predictive patterns. Most classical techniques rely on transforming raw data into structured feature vectors that capture statistical or temporal properties. Feature engineering typically includes statistical features (means, variances, percentiles, and higher-order moments), temporal features (sliding-window summaries, signal slopes, and autocorrelation), frequency features (spectral power, frequency peaks, and transform-based coefficients), and domain-specific descriptors (features derived from prior knowledge). Feature normalization, dimensionality reduction, and correlation analysis are often applied to prepare datasets for efficient learning. An overview of the pipeline is shown in Figure 2.10.

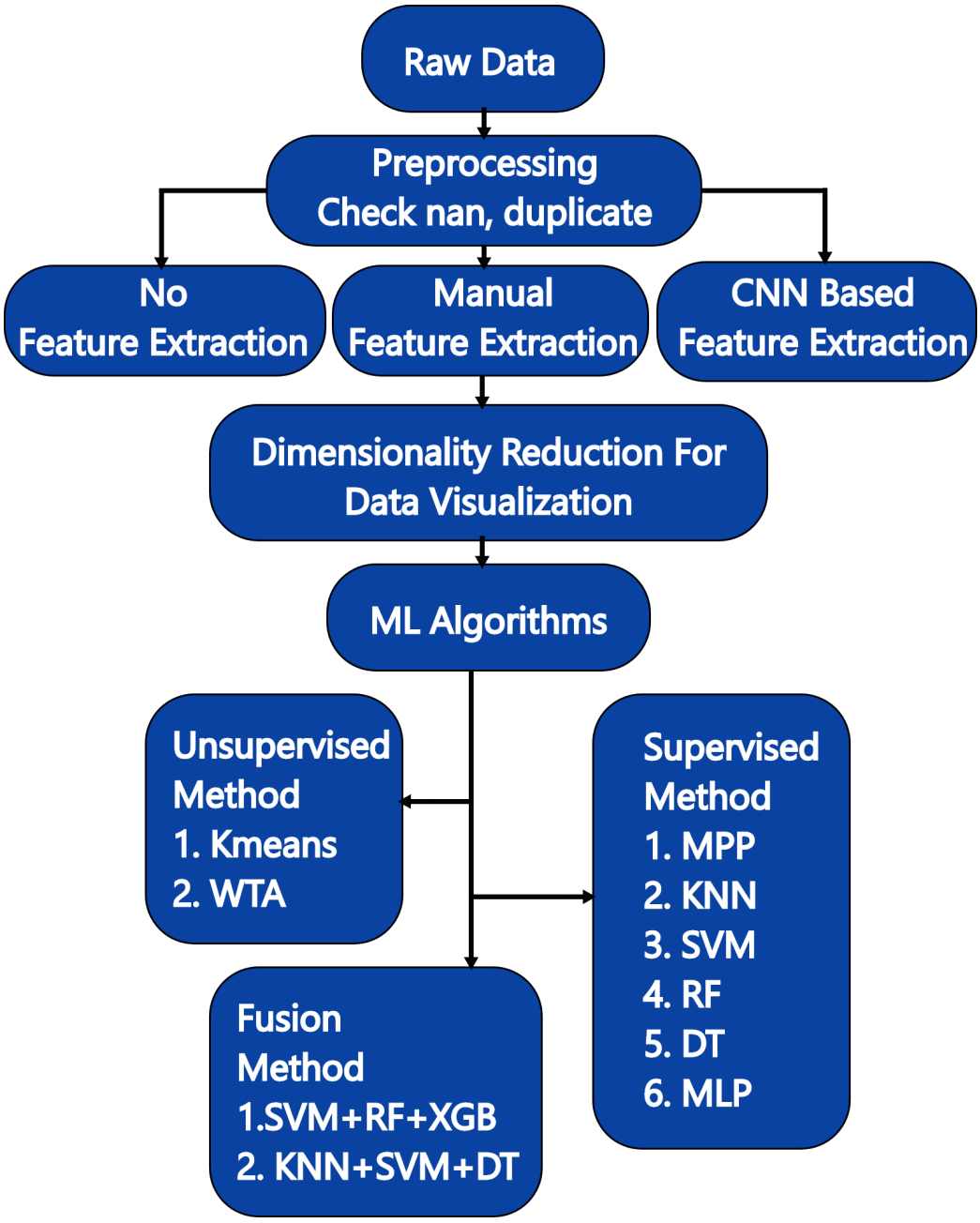


Figure 2.10: A Generic Classical Machine Learning Pipeline. (WTA: Winner-Take-All, SVM: Support Vector Machines, RF: Random Forest, XGB: X-Gradient Boosting, DT: Decision Tree, MPP: Maximum Posterior Probability, KNN: K-Nearest Neighbor, MLP: Multi-Layered Perceptron)

2.3.1 Classical Models for Classification

Classification involves assigning discrete labels to input samples based on previously learned decision boundaries. Various widely used classical models include:

- **Logistic Regression**: A linear classifier that models class probabilities using the logistic function.
- **Support Vector Machines (SVM)**: Constructs maximum-margin hyperplanes to separate classes. Kernel methods help create nonlinear decision boundaries.
- **Random Forest**: An ensemble of decision trees that averages their outputs to reduce overfitting and improve robustness.
- **Gradient Boosted Trees**: Iteratively improves weak learners using gradient-based boosting and can often achieve better classical performance.
- **K-Nearest Neighbors (KNN)**: A non-parametric, distance-based method that classifies based on similarity in feature space.

Model selection and hyperparameter tuning, kernel selection, regularization strength, and number of trees *etc.* are critical for optimizing classifier performance. Performance is typically evaluated using accuracy, precision, recall, F1-score, and confusion matrices.

2.3.2 Regression Methods

Regression techniques work with continuous-valued outputs and are critical for prediction tasks involving real-valued quantities. Common classical regression models include:

- **Linear Regression**: Models the relationship between inputs and outputs with a linear function and is useful as a baseline model.

- **Support Vector Regression (SVR)**: Applies SVM principles to regression.
- **Decision Tree and Random Forest Regression**: Captures nonlinear relations through hierarchical splits and ensemble averaging.
- **Gradient Boosted Regression**: Builds strong predictive models through sequential error minimization.

Residual analysis, mean squared error (MSE), root mean squared error (RMSE), and mean absolute error (MAE) are standard metrics for evaluating regression performance.

2.3.3 Time-Series and Sequential Prediction Models

Many practical problems involve forecasting future values based on historical observations. Classical and neural sequence models support time series reasoning through these 3 mostly used algorithms.

- **Recurrent Neural Networks (RNNs)**: Maintain a hidden state capturing past context to map time series dependencies.
- **Long Short-Term Memory Networks (LSTM)**: Uses a gating mechanism that enables the preservation of long-range information. Also addresses vanishing gradient issues of standard RNNs.
- **Bidirectional LSTM (BiLSTM)**: Processes sequences in both forward and reverse directions to encode better context.

These models are trained on sliding windows or segmented sequences and are commonly optimized using gradient-based methods. Forecasting performance is often measured using prediction error curves, or deviations from ground truth trajectories.

2.4 Detection with Quantum Machine Learning

Quantum computing can be leveraged to address computationally demanding tasks involved in detecting anomalies. Quantum Machine Learning (QML) refers to machine learning models that incorporate quantum computation-either fully quantum or hybrid quantum-classical architectures [50]. These include methods such as quantum support vector machines [51], quantum neural networks [52], and related models designed to identify deviations from expected system behavior. Such deviations or *anomalies* are often signal-system malfunctions or data tampering. They are common indicators of cyberattacks in cyber-physical control systems. Quantum models can capture intricate correlations and high-order patterns that may be difficult for classical methods to detect, making them promising candidates for anomaly detection in complex environments.

2.4.1 Quantum Computing

The remainder of this chapter introduces the essential quantum computing concepts needed to understand our quantum machine learning framework.

Classical computing represents information using binary digits, or *bits*, which exist strictly in the states 0 or 1. This deterministic representation supports all modern electronic computation but also imposes limits on scalability and parallelism. Quantum computing offers an alternative paradigm grounded in the principles of quantum mechanics. Its fundamental unit of information, the *qubit*, behaves as a two-state quantum system that can exist in a superposition of basis states. A qubit can be represented in Dirac notation as follows:

$$|\psi\rangle = c_0|0\rangle + c_1|1\rangle \tag{2.2}$$

where $c_0, c_1 \in \mathbb{C}$ and $|c_0|^2 + |c_1|^2 = 1$. Qubits may be physically realized through technologies such as superconducting devices [53] and trapped ions [54].

When multiple qubits are combined, the joint state spans a 2^N dimensional space. It allows simultaneous representation of exponentially many configurations through superposition. These quantum properties enable powerful capabilities, such as large-scale linear algebra operations and certain NP-hard computations, that are difficult for classical architectures to emulate efficiently.

2.4.2 Quantum Gates

Quantum operations are implemented through *unitary* transformations known as quantum gates. Gates may act on one or several qubits and can modify probability amplitudes or generate *entanglement*, which is a non-classical correlation between qubits.

Controlled-NOT Gate (C-NOT)

The Controlled-NOT (C-NOT) gate, shown in Figure 2.11c, is a two-qubit gate used extensively to generate entanglement. The target qubit flips its state when the control qubit is $|1\rangle$. Its matrix representation is

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix} \tag{2.3}$$

Hadamard Gate (H)

The Hadamard gate (H), depicted in Figure 2.11a, is a single-qubit gate commonly used to create equal superpositions of $|0\rangle$ and $|1\rangle$. It is defined as

$$H = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 0 \\ 1 & -1 \end{bmatrix} \tag{2.4}$$

Rotation Gates

Rotation gates (R_x, R_y, R_z) are parametrized single-qubit gates that rotate quantum states on the Bloch sphere (Figure 2.12). A general single-qubit state can be expressed as

$$|\psi\rangle = \cos\theta|0\rangle + \sin\theta e^{i\varphi}|1\rangle \tag{2.5}$$

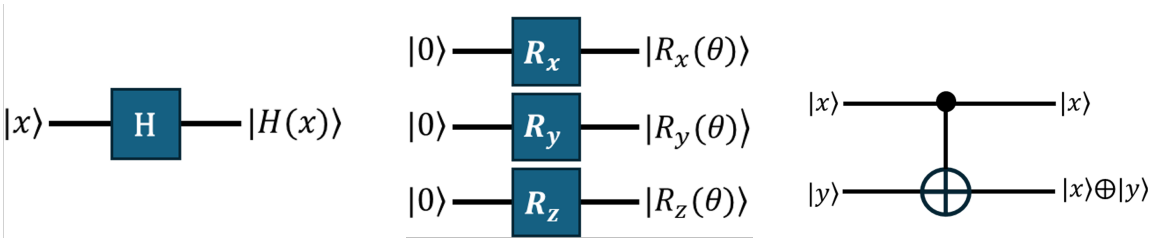
where θ and φ determine its orientation on the sphere. These parameterized rotations are widely used to encode classical data in quantum circuits. The notation for these gates is shown in Figure 2.11b.

2.4.3 Embedding Classical Data to Quantum

Classical data values like sensor measurements, are incorporated into quantum circuits through *encoding* schemes. Angular encoding is widely used due to its simplicity, which uses parametrized rotation gates to map features to superposed qubit states. An extended angle encoding approach that stores two classical features per qubit using θ and ϕ rotations [55]. In general, we denote this encoding process as a unitary transformation S_x that maps a feature vector $x = [x_1, \ldots, x_N]^T \in \mathbb{R}^N$ to an $N/2$ qubit quantum state.

2.4.4 Noisy Intermediate-Scale Quantum (NISQ) Computers

Despite rapid progress, practical quantum computing faces significant challenges due to *decoherence*, where interactions with the environment degrade quantum states. As long as such errors continue to exist, results from quantum algorithms cannot be considered strictly deterministic [56]. Modern devices fall under the category of Noisy Intermediate-Scale Quantum (NISQ) computers, which feature modest qubit counts and limited coherence times. Contemporary hardware continues to reduce error rates. For instance, IonQ's Forte system exhibits an error-per-layered-gate (EPLG)



(a) Hadamard Gate (H) (b) Rotation Gates (R_*) (c) Controlled-NOT Gate (C-NOT)

Figure 2.11: Quantum Gate Symbols

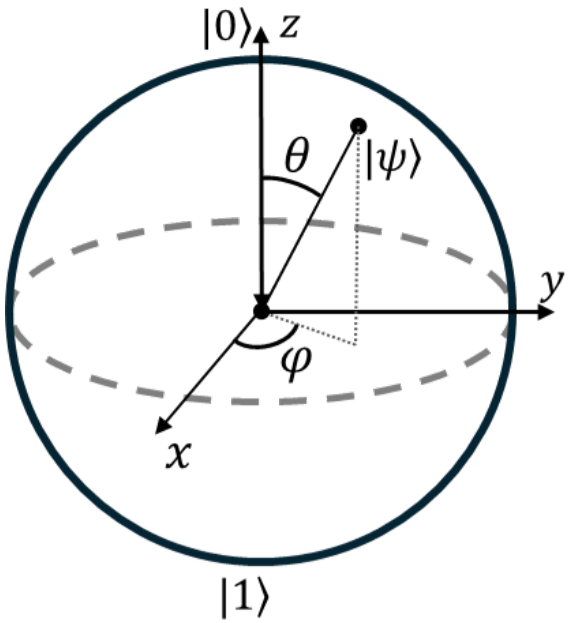


Figure 2.12: Bloch Sphere Showing a State $|\psi\rangle$ with Parameterized θ and φ

of approximately 0.4% [54], while IBM's Heron R2 achieves around 0.348% [57]. These advances are supported by emerging quantum error-correction techniques [58]. Fortunately, many machine learning tasks are tolerant to noise and can operate with approximate solutions, as demonstrated by Clark *et al.* [59].

2.4.5 Quantum Benefits

Quantum machine learning offers several advantages when dealing with complex, high-dimensional, or weakly structured data. By exploiting superposition and entanglement, quantum models can embed data into exponentially large feature spaces using limited physical resources. This enables richer feature representations and can accelerate optimization and pattern recognition tasks. Hybrid quantum–classical approaches further reduce model complexity by pairing classical optimization with quantum feature extraction that allows efficient training and maintains strong performance. These properties make QML well suited for applications such as biomedical signal analysis, cryptography, and financial modeling. As quantum hardware continues to advance, QML is expected to handle tasks that are infeasible for classical algorithms.

2.4.6 Quantum-Hybrid Machine Learning Framework

Quantum hybrid machine learning integrates classical and quantum computation within one framework. Classical preprocessing is used to normalize data and encode it into quantum states through amplitude, angle, or other encoding strategies. A parameterized quantum circuit then transforms these states and generates features with complex nonlinear correlations. After measurement, the resulting quantum outputs are passed to classical components for further processing and parameter updates. This workflow, illustrated in Figure 2.13, reduces the depth of quantum circuits and is compatible with current noisy intermediate-scale quantum (NISQ)

devices. Common hybrid approaches include quantum hybrid support vector machines, quantum hybrid neural networks, and quantum K-nearest neighbors.

2.4.7 Quantum Fidelity Kernels

Quantum-Hybrid Support Vector machines rely on quantum kernel functions derived from the fidelity, or overlap, between two quantum states produced by a parameterized feature map $U(x)$. The fidelity kernel is defined in Eq. 2.6 and serves as a quantum analog of the classical inner product. Quantum feature maps [60] encode data features into an n-qubit Hilbert space, and additional gates such as C-NOT operations introduce entanglement through the unitary $G(x)$. Starting from a known basis state such as $|0\rangle^{\otimes n}$, the feature map produces a state $\phi(x)$ whose fidelity with another state forms the kernel:

$$K(x_i, x_j) = |\langle \phi(x_i) | \phi(x_j) \rangle|^2 \tag{2.6}$$

Efficient computation of state fidelity using n qubits is based on the Hilbert–Schmidt inner product, as described by Havlicek *et al.* [61]. Figure 2.14 illustrates this process: the feature map unitary $U(x_i)$ is applied, followed by its inverse $U^\dagger(x_j)$, and the system is measured. The probability of returning to the base state $|0\rangle$ provides the fidelity between x_i and x_j. Repeating this for all data pairs yields a kernel matrix with values in $[0, 1]$, which can then be used within a one-class SVM for unsupervised anomaly detection.

2.4.8 Quantum-Hybrid Support Vector Machine

The One-Class Support Vector Machine (OCSVM) is a widely used method for anomaly detection, originally derived from the classical SVM framework proposed by Cortes and Vapnik [62, 63]. OCSVM is particularly effective when the training data consists almost entirely of normal observations with few or no anomalous examples. The model learns a boundary surrounding the majority of data points and identifies

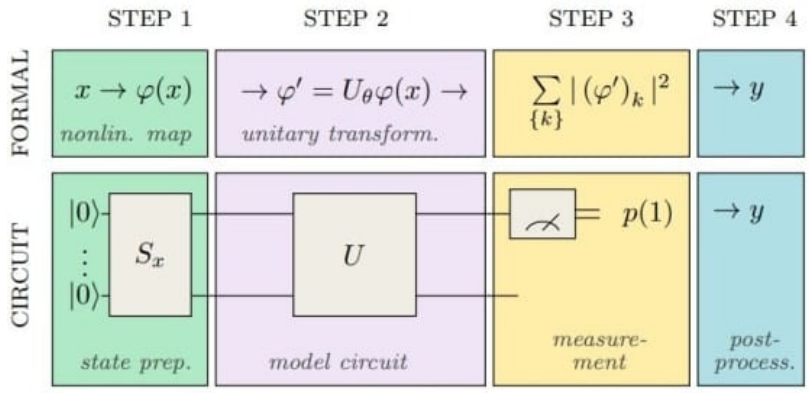


Figure 2.13: Quantum Hybrid Machine Learning Framework [6]

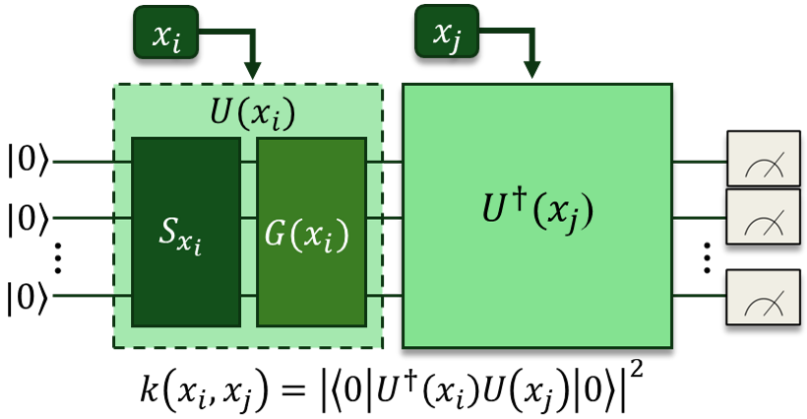


Figure 2.14: Fidelity Quantum Kernel $k(x_i, x_j) = |\langle 0|U^\dagger(x_i)U(x_j)|0\rangle|^2$. (The vectors x_i and x_j are N-dimensional features encoded into quantum states using S_{x_i}, with optional entangling gates $G(x_i)$. The circuit applies $U(x_i)$, followed by its inverse $U^\dagger(x_j)$, and measures the overlap to compute the kernel)

anomalies as points lying outside this region. Kernel functions map inputs to a high-dimensional feature space that helps draw nonlinear boundaries.

A recent extension of this approach incorporates quantum kernels [61]. In a Quantum-Hybrid SVM (QSVM), the classical kernel is replaced with the quantum fidelity kernel described earlier. The resulting quantum kernel matrix captures high-dimensional relationships encoded by the quantum circuit. This provides a powerful representation for downstream anomaly detection. Because QSVMs project data through a quantum feature map before returning to classical computation, they can identify subtle deviations that may not be captured by classical kernels.

Chapter 3

Utilizing Machine Learning for Context-Aware Digital Biomarker of Stress in Older Adults

This chapter presents a classical machine learning framework for identifying digital biomarkers of stress using multimodal physiological data collected from wearable devices. The focus is on characterizing how distinct sensor-derived patterns, combined with contextual task information from stress-induction protocols, can serve as reliable indicators of stress. Incorporating context such as task phase and cognitive load allows the modeling process to distinguish between normal physiological variability and genuine stress responses. These ideas are important because physiological expression of stress differs across individuals, and contextual grounding helps reduce ambiguity when interpreting sensor-derived biomarkers.

Beyond detection, this chapter also develops predictive models that estimate short-term future stress states by leveraging the dynamics in physiological signals. Wearable sensors record real-time variations in electrodermal activity, blood volume pulse, acceleration, and skin temperature, all of which reflect physiological changes associated with stress. These and are highly relevant for older adults whose

physiological patterns may differ due to aging complications or medication effects. By combining context-aware classification with time-series prediction, this chapter establishes a core set of digital biomarkers that represent both the presence of stress and its probable progression. These biomarkers form the foundational baseline on which the deep learning and quantum machine learning models in later chapters are built.

3.1 Related Works and Motivation

In recent years, the advancement of machine learning has led to the gradual integration of deep networks into the realms of stress detection and emotion recognition based on digital biomarkers. Digital biomarkers not only enable accurate stress monitoring but also pave the way for personalized interventions. With continuous data collection, algorithms can learn individual patterns and preferences, tailoring interventions to suit each person's unique needs. This personalized approach has the potential to enhance the effectiveness of stress management strategies, moving away from generic solutions to more targeted and impactful interventions.

For instance, Saylam *et al.* [33] conducted a study on stress biomarkers, aiming to predict stress levels by identifying crucial parameters from various modalities, including mobile phones and wearables. They utilized a ranking system to emphasize the significance of different modalities, such as sleep and activity levels, in classifying stress. Daily stress responses served as labels, categorizing participants into five stress levels. The study compared the efficacy of considering only important parameters versus using all parameters, employing the Random Forest (RF) algorithm. Additionally, the researchers explored the potential impact of participants' personalities on stress levels to enhance overall understanding. Opoku *et al.* [32] conducted an analysis using a longitudinal dataset and employed statistical and machine learning methods to explore the connection between digital biomarkers, mood ratings, and depression. The study aimed to determine if there

are differences in digital biomarkers and mood ratings between depressed and non-depressed participants. The findings revealed that accurately predicting the depression status of participants is possible using either only digital biomarkers or a combination of digital biomarkers and mood ratings as predictors.

Neural networks have also been employed in this field in recent times. By converting multi-modal sequence signals into pictures and using an integrated, scalable, low-power Deep Convolutional Neural Network (DCNN) to learn common features, Jafari *et al.* [64] achieved an impressive 94% accuracy rate in stress detection. In order to extract features from physiological data and behaviors gathered by wearable sensors, Aristizabal *et al.* [65] used deep networks. In order to extract deep features from EDA, PPG, and Zygomaticus Electromyography (zEMG), Hassan *et al.* [66] used the Deep Belief Network. These characteristics were then merged with statistical features. In order to extract their image features, Siddharth *et al.* [67] used deep networks to create spectrograms from every signal channel.

Despite the remarkable performance improvements achieved by deep learning-based algorithms in stress assessment, they come with only a few dozen subjects and struggle to meet the requirements necessary for effectively training deep networks. Researchers have addressed such problems and proposed solutions based on machine learning with manual statistical feature engineering and have targeted specific age groups for less spread in the training set [43, 42, 45]. These studies have relied on questionnaires or annotations by external observers for ground truth on the base level and stress level. However, such ground truths are prone to error and differ from person to person. Hence, Nath *et al.* [68] used cortisol as a stress ground truth biomarker and also incorporated context as an added input. This concentration of stress hormone levels ensures unbiased stress labeling for that specific age group. Table 3.1 shows an overview of recent works that deal with stress detection for older adults and take context information into account.

The majority of currently released works, however, are focused on younger adults or fail to take advantage of relevant contexts. However, literature in this domain shows

Table 3.1: Summary of Recent Research on Older Adult Stress Detection

Authors & Year	**Signals**	**Feature Extraction**	**Base Model**	**Used Context**	**Stress Ground Truth**
Ferreira *et al.* [46] 2014	ECG, EEG, EDA	Manual	Quadratic Discriminant Analysis	✗	Questionnaire
Kikhia *et al.* [43] 2016	EDA	Manual	Thresholding	✗	Annotation by clinical Staff
Belk *et al.* [48] 2016	EDA, IMU, HR, ST, Grip Force	Manual	Bayesian Probability	✗	Annotation by external observer
Delmastro *et al.* [42] 2020	ECG, EDA	Manual	Support Vector Machine,K-NN, Decision Tree	✗	Predetermined annotation
Cheong *et al.* [45] 2020	ST, HR, SC, HUM, AT	Manual	Statistical Correlation	✗	Questionnaire
Nath *et al.* [68] 2021	EDA, BVP, IBI, ST	Manual	Decision Tree	✗	Cortisol Concentration
Proposed Work 2023	**EDA, BVP, IBI, ST**	**Manual** **Automatic**	**Random Forest** **1-D CNN**	✓ ✓	**Cortisol Concentration**

ECG: Electrocardiogram, EEG: Electroencephalography, EDA: Electrodermal Activity, SCL: Skin Conductance Level, IMU: Inertial Measurement Unit, HR: Heart Rate, ST: Skin Temperature, SC: Step Counter, HUM: Humidity, AT: Air Temperature, BVP: Blood Volume Pressure, IBI: Inter Beat Interval

that biomarkers like oxidative stress, cortisol, and metabolic imbalances are highly correlated with age and are fairly distinguishable among the age groups [69]. Similarly, the use of cortisol concentration as a stress biomarker has also been investigated. To calculate stress levels, the majority of studies have relied on subjective procedures or surveys. While there have been recent experiments in context-based stress detection, feature selection requires significant human involvement. The next step toward improvement will require an automatic feature encoder that employs context information as well as cortisol as ground truth. With that in mind, we incorporated a 1-D CNN-based feature encoder. CNNs are a type of artificial neural network (ANN) with convolutional layers in its hidden layers [70]. It was first used in computer vision tasks, but it has since piqued the interest of researchers in a wide range of fields, including bio-signal classification such as electromyogram (EMG) signal classification for gesture recognition and electroencephalogram (EEG) pattern identification for assistive machine control [71, 72, 73, 74]. CNN has been used effectively in arrhythmia detection [74], signal component identification [75], biometric recognition [76], and other applications in ECG signal classification. These experiments revealed the power of CNN in biosignal categorization, demonstrating that CNN has the capacity to identify stress. As a result, there is an immediate need for an end-to-end CNN-based context-aware machine learning model that uses salivary cortisol as a stress biomarker for stress detection in older persons.

While various approaches have been presented for detection [38, 36, 29, 35, 77, 78, 43, 48], there is still room for improvement in forecasting. A variety of stress-monitoring and management technologies have been proposed over the years. For example, a clinical decision support system was described in [79], which used case-based reasoning, textual information retrieval, rule-based reasoning, and fuzzy logic techniques with finger temperature to monitor stress. The results show that the algorithm is valid when compared to human expert assessments.

In [80], the authors present a driving stress monitoring system that combines correlation analysis with a mathematical stress estimation function. They used

Table 3.2: Overview of Recent Literature on Stress Prediction

Authors & Year	**Sensors**	**Target Age Group**	**Ground Truth**
Jaques *et al.* [25] 2015	EDA, ACC, ST	Young Adults	Daily Survey
Jaques *et al.* [24] 2015	ACC, SCL	Young Adults	Self Reporting
Taylor *et al.* [23] 2017	ACC, SCL	Young Adults	Self Reporting
Reddy *et al.* [82] 2018	ECG	Young Adults	Questionnaire
Alharthi *et al.* [83] 2019	ECG	Older Adults	Daily Survey
Parent *et al.* [84] 2019	EDA, IBI, BVP, BR, SCL	Young Adults	Types of Game Played
Yu *et al.* [85] 2019	ACC, ST, EDA	Young Adults	Self Reporting
Martino *et al.* [86] 2020	EDA, ECG, HR	Young Adults	Questionnaire
Umematsu *et al.* [87] 2020	EDA, IBI, BVP, BR, SCL	Young and Older Adults	Self Reporting
Dai *et al.* [88] 2021	ECG, EDA	Young and Older Adults	Self Reporting
Kene *et al.* [89] 2021	ECG	Young and Older Adults	Questionnaire
Proposed Work 2024	**EDA, BVP, ST**	**Older Adults (60-80) years**	Cortisol



Figure 3.1: Empatica Wristband for Data Collection

physiological data like ECG, electromyography (EMG), respiration rate (RR), and galvanic skin response (GSR) to extract characteristics for stress classification. In addition to stress monitoring and management, some of these systems provide users with real-time interventions during stressful events and activities. MoodWings [81] is a wearable biofeedback gadget designed to help drivers calm down and relax during stressful driving circumstances. The device uses *Ambu Blue Sensor gel* sensors to collect EDA and ECG readings. The technology uses a butterfly wing motion as a simple feedback mechanism to notify drivers of their stress levels.

In the StudentLife study [90], the relationship between passive sensor data and students' psychological well-being was investigated using standard mental health surveys. The study found that interpersonal interactions and academic college life have a considerable impact on students' stress levels. Several research [91, 92, 93, 94, 88] attempted to link passive mobile phone sensing data to mental health issues. These studies sought to uncover relationships between stress and time-varying variables, as well as forecast stress levels.

Thelwall [95] introduced *TensiStrength*, a system that uses a rule-based lexical approach to detect stress and anxiety. Lin *et al.* [96] presented a factor graph model linked with a CNN for stress detection, with fairly comparable results obtained by training the model on Twitter and SinaWeibo data. Pankajavalli *et al.* [97] proposed an Adaptive Neuro-Fuzzy Inference System that uses the FireWorks GreyWolf Optimization (ANFIS-FWGWO) classification algorithm to forecast stress from a sensor-equipped keyboard. Reddy *et al.* [82] trained their model using a variety of machine learning approaches after thoroughly cleaning and preparing the data. Boosting algorithms outperformed the other models in terms of accuracy. Stutz *et al.* [98] sought to examine whether smartphone usage and sensor data collected automatically could predict a user's perceived stress levels with a customized abbreviated version of the Perceived Stress Scale (PSS). To do this, they conducted user research in which smartphone data and stress levels were collected over two weeks. Their findings showed substantial relationships between stress scores, smartphone

usage, and sensor data that provided novel approaches to autonomously monitor stress using smartphone technology. Sharma *et al.* [99] evaluated the efficacy of machine learning algorithms in predicting stress levels among undergraduate students to promote early intervention and treatment. The dataset, which included data from over 200 students collected using the PSS scale, was obtained from the university. Several classification methods were used, and their investigation revealed that the Random Forest classifier has the greatest accuracy of 94.73%.

Kene *et al.* [89] explored multiple machine learning methods on the *Physiobank* dataset to predict stress levels. They achieved around 83% F1-score with gradient boosting algorithms. Recently, Ding *et al.* [100] presented an effective way for precisely and effectively classifying stress into several target classes: low-normal, medium-low, medium, medium-high, and high. The suggested method uses a hybrid model (HB) that combines a gradient boosting machine (GBM) and random forest (RF). These models are integrated using soft voting criteria, with each model's prediction probability contributing to the overall prediction. Later, Haque *et al.* [101] reviewed sensing technologies and preprocessing strategies for stress prediction with multi-modal data. This is followed by a thorough examination of how a number of Machine Learning (ML) models have been used to predict stress using Heart Rate Variability (HRV) measurements.

Table 3.2 summarizes that existing studies still show gaps for improvement, mostly due to the restricted evaluation of stress in conjunction with other human factors as well as variations in stress assessment methods [102, 103]. From the targeted age group and the ground truth labeling technique, it is observed that there is an under-explored domain for stress prediction in older adults with cortisol biomarkers.

3.2 Contributions

In this work, we first test the performance of *Random Forest,* a machine-learning model for stress detection, with cortisol concentration as ground truth. We also

incorporate contextual information recorded during the test protocol. After the trial with manual feature engineering, we also automate the feature selection technique with a convolutional neural network (CNN) based feature encoder. Finally, the chapter ends with a framework that predicts future stress episodes based on past data.

Here are the highlights of our contributions:

- We explore the effectiveness of machine-learning models to find the correlation of digital biomarkers of stress with experimental data from 40 healthy older adults.

- We propose a ground truth labeling scheme based on cortisol concentration. We labeled stress into three distinct classes: *No-Stress, Low-Stress,* and *High-Stress*.

- We investigate the efficacy of digital biomarkers from three signal streams (EDA, BVP, and IBI) for stress classification. We also validated that the fusion of sensors enhances the accuracy of the machine learning classifier when compared to the case of a single signal stream.

- We also propose a CNN-based feature encoder that automates the feature selection process and selects the best possible inputs for the FCN.

- We finally report an about 4% increase in accuracy and a 0.3 increase in the F-1 score for the CNN-based feature extraction compared to the stress detection method with Random Forest on our dataset.

- We also provided a framework with two time series-based machine learning models, RNN and LSTM, to predict stress based on trends in past data and investigated the effectiveness of the proposed framework on our dataset.

The data recording with the TSST protocol is discussed in Chapter 2 Section 2.1 Subsection 2.1.1. The rest of this chapter is organized as follows: The classical method for stress detection is presented in Section 3.4 and the CNN based detection method is

discussed in Section 3.5. Finally, the prediction framework is presented in Section 3.6. Section 3.7 presents the conclusion with future direction.

3.3 Data Collection, Preprocessing and Labeling

In this section, we will first explain the procedure of our data collection from the selected participants. Then we mention the stress protocol that was incorporated to induce stress in participants. Then we provide our ground truth labelling technique and its visualization after labelling. Afterwards, we explain the incorporation of contextual information, followed by feature selection and preprocessing.

3.3.1 Data Collection

In our investigation, a group comprising 40 elderly individuals ranging in age from 60 to 80 years was selected as the subject population. Within this sample, the demographic distribution is comprised of 28 female and 12 male participants. Due to data corruption, the dataset of one participant was deemed unsuitable for inclusion, resulting in a final dataset derived from 39 participants. To provide an unbiased dataset, a thorough screening procedure was conducted on each participant prior to their registration in the experimental protocol to determine whether they had any pre-existing medical illnesses or disorders. We collected physiological data using the Empatica E4 wristband, which has built-in sensors for skin temperature (ST), photoplethysmography (PPG), and electrodermal activity (EDA), as illustrated in Figure 3.1.

3.4 Context-Aware Stress Detection with Supervised Machine Learning

Our method for context-aware stress detection with ML consists of 3 main parts. First, we have the statistical feature extraction. After collecting the data from the Empatica wristband, they go through preprocessing steps where they are cleaned, filtered, normalized, and fused. This sensor fusion plays an important role in removing false positives, which we will show in more detail in the later part of the section. Second, the ground truth is estimated from the cortisol concentration, and the context information comes from the TSST protocol. Finally, they are all fed to the *Random Forest* model for training. The overall workflow illustrated in Figure 3.2 is explained in this section.

3.4.1 Statistical Feature Extraction from Signal stream

After collecting sensor data, they are preprocessed for the corresponding machine-learning algorithm. The raw Electro-Dermal Activity (EDA) and Blood Volume Pressure (BVP) signals are normalized and filtered using a 5th-order low-pass filter with cut-off frequencies of 1 Hz for EDA and 10 Hz for BVP. The Inter Beat Interval (IBI) signal undergoes no preprocessing. Subsequent to preprocessing, features are extracted using a window of 1.5 minutes with an overlap of 0.75 minutes from the samples. Previous studies have indicated that characterizing EDA peaks or startles can provide insights into an individual's stress level [104]. The first derivative of the EDA signal undergoes a peak detection algorithm [105], and features such as peak amplitude, peak width, and peak prominence are computed. In total, 18 characteristics are extracted from the EDA signal. The BVP signal, obtained from the photoplethysmogram (PPG) sensor, reflects heart activity. The cardiovascular arousal associated with increased blood pressure and heart rate during stress is correlated with the BVP signal [106]. Seventeen characteristics are derived from the

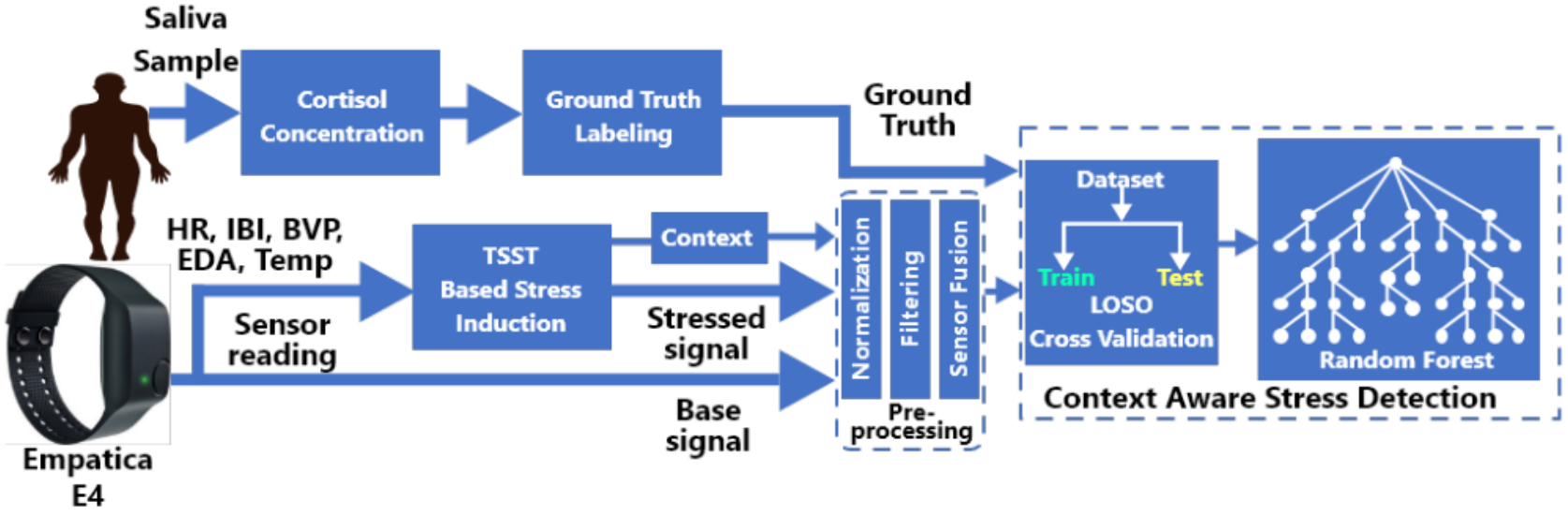


Figure 3.2: Workflow of Stress Detection with Machine Learning Model(Random Forest)

BVP signal, summarizing peak amplitude, breadth, and prominence similar to EDA features. Additionally, the IBI and ST signals contribute six time-domain statistical characteristics each. The IBI, representing the interval between two heartbeats, is instrumental in understanding cardiac activity concerning stress. ST refers to the skin temperature in that instance.

3.4.2 Machine Learning Model

We have chosen *Random Forest* as our base model for the classification problem. Supervised machine learning algorithms, such as Random Forest, are commonly utilized for addressing classification and regression challenges. It constructs a group of decision trees during training and outputs the class that is the mode of the classes (classification) or mean prediction (regression) of the individual trees. The algorithm is as follows:

1. **Bootstrap Sampling:** The dataset is broken into subsets to send randomly in individual trees.

2. **Decision Trees:** For each sample and feature subset, a decision tree is constructed. The splitting is based on a preset criterion. For the classification problem, we used logarithmic Gini impurity shown in Eq. (3.1)

$$Gini(D) = 1 - \sum_{i=1}^{C} (P_i)^2 \tag{3.1}$$

   Here, D is the dataset at that tree node, C is the number of classes iterated by i and P_i is the probability of class i in node D.

3. **Ensembling and Voting:** Based on the criterion, the required number of trees is constructed. The loss is calculated with logarithmic entropy shown in Eq. (3.2). These ensemble trees will provide a prediction of their own and out of it, the majority-voted class will be the final prediction. The equation for the

Table 3.3: Training Parameters for Random Forest

Parameter		**Value**
No of Estimators	-	40
Criterion	-	Entropy
Minimum Samples Split	-	2
Maximum Depth	-	Till (minimum sample Split - 1)
Minimum Samples Leaf	-	1
Maximum Feature	-	Auto
Bootstrap	-	True
Random State	-	4

Table 3.4: Extracted Statistical Features from Physiological Signal

Feature	**Description**	**Equation**
Mean	The average value of the signal	$x_\mu = \frac{1}{N}\sum_{i=1}^{N} x_i$
Median	The middle value of the signal when sorted	Median $= \frac{1}{2}(x_{\frac{N+1}{2}} + x_{\frac{N+2}{2}})$
Mode	The most frequently occurring value in the signal	–
Standard Deviation	A measure of dispersion around the mean	$x_{sd} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i - x_\mu)^2}$
Variance	The average of squared differences from the mean	$x_{var} = \frac{1}{N}\sum_{i=1}^{N}(x_i - x_\mu)^2$
Skewness	A measure of asymmetry in the signal's distribution	–
Kurtosis	A measure of sharpness of the peak	–
Range	The difference between maximum and minimum values	Range $= \max(x) - \min(x)$
Interquartile Range (IQR)	Range between the first and third quartiles	IQR $= Q_3 - Q_1$
Root Mean Square (RMS)	Square root of the average of squared values	RMS $= \sqrt{\frac{1}{N}\sum_{i=1}^{N} x_i^2}$
Coefficient of Variation (CV)	Standard deviation to mean ratio	CV $= \left(\frac{x_{sd}}{x\mu}\right) \times 100$
Entropy	A measure of signal disorder or unpredictability	–
Autocorrelation	Correlation of signal with itself at different time lags	–
Zero Crossing Rate	Rate at which the signal changes sign	–
Time of Max and Min Values	Points in time when signal reaches max and min values	–
Signal Energy	Total energy contained in the signal	Energy $= \sum_{i=1}^{N} x_i^2$
Signal Amplitude	Magnitude of the signal	Amplitude $= \max(\lvert x_i \rvert)$
Waveform Length	Cumulative length of the signal waveform	Wave Length $= \sum_{i=2}^{N} \lvert x_i - x_{i-1} \rvert$

prediction of individual trees is shown in Eq. (3.3) and the final prediction is shown in Eq. (3.4). Here $\hat{Pred}_i(x)$ denotes the prediction of a tree, $f(.)$ is the indicator function, and $L_i(x)$ is the leaf node to which x is assigned in tree i.

$$Entropy = -\sum_{i=1}^{C} P_i \times Log_2(P_i) \tag{3.2}$$

$$\hat{Pred}_i(x) = \text{argmax}_c \left(\sum_{x_j \in L_i(x)} f\left(y_j = c\right) \right) \tag{3.3}$$

$$Pred = argmax_i(\hat{Pred}_i) \tag{3.4}$$

3.4.3 Model Training and Testing

After extracting features from the data, we employed salivary cortisol analysis to label these features as indicative of 3 stress classes. To evaluate the performance, we utilized a *Leave One Sample Out (LOSO)* cross-validation technique. This involved iteratively using one participant's data for testing while training on the remaining participants' data. During the training phase, a feature selection process based on Kendall's tau correlation was implemented to identify the most pertinent features. This correlation measure established the link between the training feature set and stress labels derived from cortisol concentration [107]. Only features exhibiting statistically significant correlations (p-values below 0.05) were retained. The final set of training features is presented in Table 3.4. When training the *Random Forest* model, hyperparameters were optimized based on the model's performance. The final selection of hyperparameters that resulted in the best accuracy on the test set is shown in Table 3.3.

3.4.4 Result Analysis

We now assess the *Random Forest* model's performance using our dataset in terms of the estimated ground truth, following training and testing. We will then talk about how well it performs with added *context*. We used the F1-score and weighted accuracy to assess the model's performance. The summarized results, as presented in Table 3.5, illustrate the performance of a single sensor and the sensor fusion. The enhancement in the F-1 score suggests that sensor fusion plays a role in mitigating the likelihood of false-positive detections. Notably, our model achieved a commendable accuracy even in the absence of context, as per the proposed ground truth labeling. Introducing context resulted in an increase of around 20% in accuracy and around 0.2 in F-1 score. The increase in F-1 score explains the model's performance in not only making the correct detection for positive cases but for negative cases as well.

3.5 Context-Aware Stress Detection with CNN-Based Automatic Feature Encoder

Our method for context-aware stress detection with a CNN-based model is based on standard, well-known CNN architectures. First, we have the automatic feature encoder, which is a series of 1-D convolutional layers. It is a sophisticated tool for extracting features and learning from 1D data sequences automatically. Based on loss, activation, and gradient values, we keep an appropriate number of neurons in each layer in our model architecture. During convolution, the CNN kernels move over the components of the 1D input signal stream. After collecting the data from the Empatica wristband, they are fed directly to CNN. Similar to ML, the ground truth is estimated from the cortisol concentration, and the context information comes from the TSST protocol. Finally, they are all fed to the feature encoder to extract a new feature representation. To test the performance for this case, we have used 5-fold

Table 3.5: Performance of *Random Forest* Model with and without Context

Criteria	Sensor list	EDA	PPG	PPG	EDA, PPG	EDA, PPG	EDA, PPG, ST
	Signal list	**EDA**	**BVP**	**IBI**	**EDA, BVP**	**EDA, BVP, IBI**	**EDA, BVP, IBI, ST**
Features	**Total**	18	17	6	35	41	47
	Selected	11	11	2	22	24	27
Without Context	**Macro F-1 Score**	0.723	0.711	0.713	0.712	0.734	0.734
	Accuracy (%)	72.95	73.40	72.77	72.51	72.44	72.44
With Context	**Macro F-1 Score**	0.922	0.907	0.910	0.909	0.937	0.943
	Accuracy (%)	93.13	93.69	92.89	92.56	92.48	91.01

cross-validation. The overall workflow illustrated in Figure 3.3 is explained in this section.

3.5.1 Automatic Feature Extraction with Feature Encoder

Our 1D CNN feature encoder consists of an *input* layer, a group of *convolutional* and *max-pooling* layer pairs, a *flattening* layer, and *dense* layers. The *convolutional* layers have a gradually increasing number of filter sizes with a ReLU activation function followed by a *max-pooling* layer. This type of pattern helps the model choose the best possible weights for the input streams. The ReLU activation function helps to avoid the vanishing gradient so that a faster convergence can be obtained, and the *max-pooling* layer has been introduced to reduce the dimensions of the feature maps. Then the *flattening* layer down-samples data into a 1D vector. Finally, there is a *dense-dropout* pair that reduces model overfitting. We have chosen the dropout rate to be 30% after trial and error for the best accuracy.

3.5.2 Fully Connected Neural Network

After the CNN-based feature encoder, the encoded features are sent to a series of *dense* and *dropout* layers for final decision-making. The dropout rate was similar to the feature encoder with a ReLU activation. Finally, the output layer sets out 3 nodes with a softmax activation function that represents three stress classes.

As the problem formulated suggests it is a multimodal classification problem, Eq. (3.7) provides the class prediction with output layer activation of *softmax* (σ_o) shown in Eq. (3.5). Other layers have activation of *relu* (σ_h) shown in Eq. (3.6). The entropy *i.e.* loss to optimise is chosen to be the *categorical cross-entropy loss* shown in Eq. (3.8).

$$\sigma_o(z) = \frac{e^{z_i}}{\sum_{j=1}^{N} e^{z_j}} \tag{3.5}$$

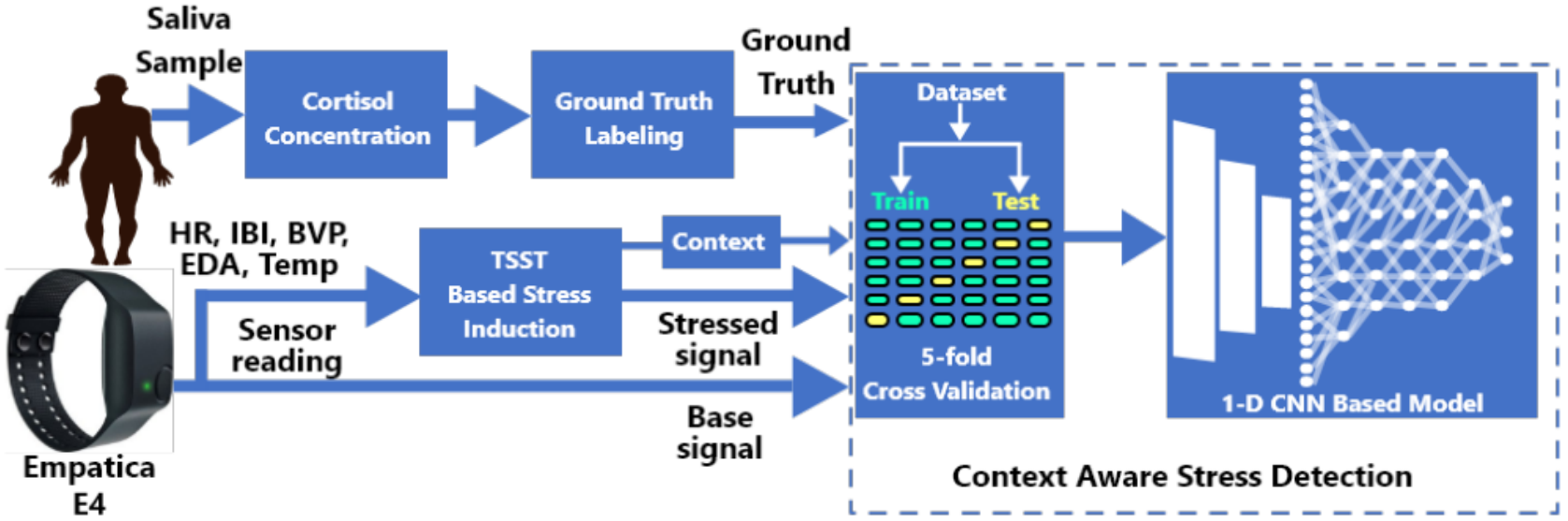


Figure 3.3: Workflow of Stress Detection with CNN-Based Model

Table 3.6: Training Parameters for Convolutional Neural Network

Parameter		**Value**
Base Architecture	-	CNN
Classes	-	3
Number of Epochs Trained	-	25
Hidden Layer Activation	-	Rectified Linear Unit (ReLU)
Output Layer Activation	-	Softmax
Optimizer	-	Adam
Loss Function	-	Categorical Cross-entropy

Table 3.7: Proposed Model Layers

Model	**Layer Name**	**Layer Info**	**Number of Parameters**
Feature Encoder	Conv-1D	Filter = 32	128
	Conv-1D	Filter = 64	6208
	Maxpool-1D	–	0
	Conv-1D	Filter = 128	24704
	Maxpool-1D	–	0
	Conv-1D	Filter = 256	98560
	Maxpool-1D	–	0
	Flatten	–	0
	Dense	Node = 2048	3147776
	Dropout	Rate = 30%	0
	Dense	Node = 512	1049088
	Dropout	Rate = 30%	0
Fully Connected Network	Dense	Node = 128	65664
	Dense	Node = 32	4128
	Dense	Node = 8	264
	Dense	Node = 3	27
Total Parameters			4,396,547

$$\sigma_h(z) = max(0, z) \tag{3.6}$$

$$class = arg_max \left[\sigma \left\{(W_i)^T \times \phi(X_j) + b^i\right\}\right] \tag{3.7}$$

$$Loss(y, \hat{y}) = -\sum_{i=1}^{N} y_i \cdot \log(\hat{y}_i) \tag{3.8}$$

Here, W_i, X_j, b_i are the weight, input vector, and bias for i^{th} and j^{th} instances respectively. N is the total number of samples, y_i is the true probability distribution (one-hot encoded) for class i, and $\hat{y}_i$ is the predicted probability for class i. The training hyper-parameters are mentioned in Table 3.6

Our proposed CNN based model consists of these two major parts cascaded together. Firstly, the feature encoder takes all possible statistical features and transforms them into a new feature representation. Secondly, these new features are fed to a fully connected neural network comprised of several dense layers. The primary determinant of a model's complexity is the quantity of trainable parameters. This comprises the total number of neurons that are employed in creating the layer-to-layer connection. With m and n serving as input and output nodes, respectively, a dense layer with a total of $(n+1) \times m$ trainable parameters is formed. Convolutional layers with input and output feature maps of p and q, respectively, combined with a filter of size $i \times j$ result in total trainable parameters of $(i \times j \times p + 1) \times q$. We present a breakdown of the number of parameters in our suggested model in Table 3.7.

3.5.3 Result Analysis

We outlined the additional advantage of sensor fusion and context information in Table 3.5. Thus, we have compiled the sensor fusion result into Table 3.10. It shows that the F-1 score increased from 0.7552 to 0.9745 and the accuracy increased from 83.7797% to 96.7525% after adding additional contextual information. The increased

F-1 score suggests that sensor fusion and added context reduce the possibility of false-positive detections. Based on the proposed neural network-based model, we obtained a reasonable accuracy without context, and after incorporating context, the accuracy increased by roughly 13%.

Cross-validation is a crucial technique in machine learning and statistical modeling that addresses the challenge of assessing a model's performance on new, unseen data. We performed 5-fold cross-validation with an 80/20 train/test split and recorded the performance in each case. From Table 3.8 it can be seen that Fold-1 has the highest average accuracy and Fold-3 has the highest macro F-1 score. On the other hand, Fold-2 has the lowest accuracy and F-1 score.

Again, for multi-class classification problems, accuracy and F-1 score might not always provide a full picture of a model's performance. This is specifically true for medical data, where false positive and false negative detection can cause serious harm. In such cases, the confusion matrix provides a quantitative measurement of the model's performance in each class. The confusion matrix generated from our experiment in Table 3.9 shows that only a few times, our model confuses *Low-Stress* with *High-Stress*. As seen from the projected dataset in Figure 2.5, this minor confusion was expected. The high value of its diagonal validates our model's performance on stress level classification.

Table 3.10 shows the performance of our proposed model with and without contextual information. With the sensor fusion, our model has an F-1 score of 0.75 and an accuracy of 83.78% when it performs without any contextual information from the TSST protocol. On the other hand, after the inclusion of contextual information, the F-1 increased to 0.97, which is 0.22 higher. Similarly, the accuracy increased by 13% to 96.75%.

Table 3.8: Performance for 5-fold Cross validation

Fold Sequence	**Avg Accuracy**	**F-1 score**
Fold-1	99.8685% (highest)	0.9888
Fold-2	86.9696% (lowest)	0.9379 (lowest)
Fold-3	99.2796%	0.9928 (highest)
Fold-4	99.3216%	0.9752
Fold-5	98.3232%	0.9778
Average	96.7525%	0.9745

Table 3.9: Confusion Matrix (weighted average of 5 folds)

Ground Truth	**Predicted Class**		
	No-Stress	Low-Stress	High-Stress
No-Stress	83.8	2.6	3
Low-Stress	5.6	243.8	2.2
High-Stress	2.6	2.8	58

Table 3.10: Performance of Proposed CNN-Based Model with and without Context

Criteria	**Sensor list**	**EDA**	**EDA, PPG**	**EDA, PPG**	**EDA, PPG, ST**
	Signal list	**EDA**	**EDA, BVP**	**EDA, BVP, IBI**	**EDA, BVP, IBI, ST**
Without Context	**Macro F-1 Score**	0.6992	0.7161	0.7492	0.7552
	Accuracy (%)	80.0078	81.4290	83.2227	83.7797
With Context	**Macro F-1 Score**	0.9022	0.9240	0.9667	0.9745
	Accuracy (%)	92.3965	94.0378	96.1092	96.7525

Table 3.11: Performance Comparison of Signal Fusion for Proposed CNN-Based Model with Machine Learning Model

Criteria		**Without Context**		**With Context**	
Sensor List	**Signal List**	**Manual Feature ML**	**CNN Based ML**	**Manual Feature ML**	**CNN Based ML**
EDA, PPG, ST	EDA, BVP, IBI, ST	f-1: 0.73 ACC: 72.44	f-1: 0.75 ACC: 83.77	f-1: 0.94 ACC: 92.48	f-1: 0.97 ACC: 96.75

3.5.4 Result Comparison and Discussion

Table 3.11 shows a comparison of performance between the machine learning model with manual feature engineering and the CNN-based model with automatic feature selection. It is seen that the CNN-based model outperforms ML models in the case of both with and without contextual information. Although the F-1 score is similar, the accuracy for the CNN-based model is around 10% higher for cases without context and 4% higher for cases with context.

3.6 Predicting Stress in Older Adults with RNN and LSTM from Time Series Sensor Data and Cortisol

Predicting stress is essential for supporting healthcare, well-being, and daily functioning. Although mild stress can be helpful, chronic stress significantly increases vulnerability to physical and psychological illness [8], including hypertension, cardiovascular disease, and ulcers [9]. It is also closely linked to depression-related indicators [10] and contributes to cognitive decline, raising Alzheimer's disease risk among older adults [11]. Accurately predicting stress levels can enable early intervention, personalized care, and better long-term health outcomes.

Stress is influenced not only by personal factors but also by environmental and organizational conditions, affecting health and productivity [108]. Chronic stress triggers harmful physiological and cognitive changes, highlighting the need for proactive monitoring. Advances in wearable sensing and artificial intelligence now make it possible to predict stress unobtrusively in real-world settings. Motivated by this potential, this work develops a framework for predicting stress in older adults, aiming to support timely detection and improve well-being.

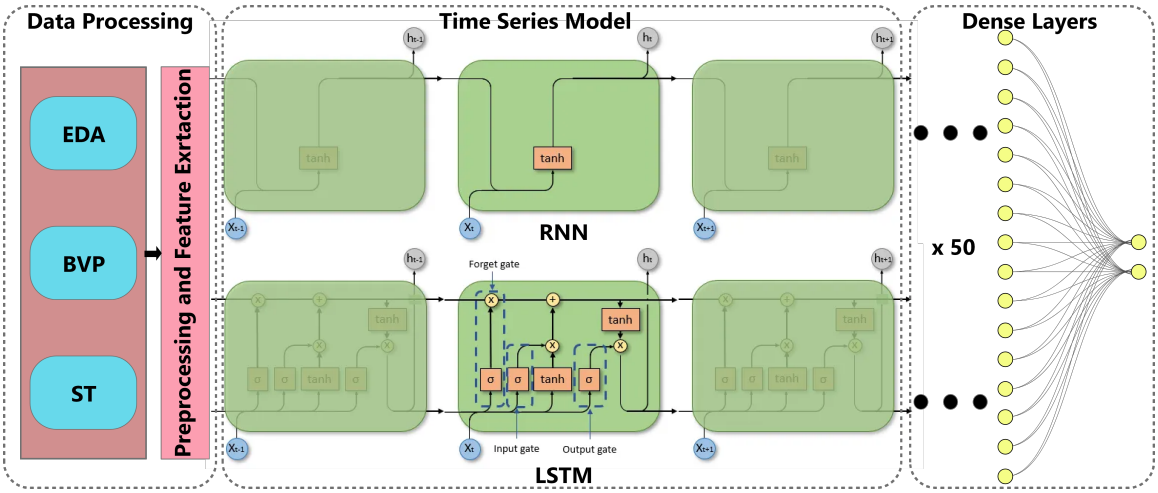


Figure 3.4: Workflow of Proposed Stress Prediction Model

Table 3.12: Training Hyper-parameter and Hardware

Parameters	Value
backbone	RNN or LSTM
Number of layers	50
classes	2(Binary)
batch size	128
Max Sequence length	$4.5min \times 60 = 270samples$
optimiser	Adam
Learning rate (Lr)	0.005
loss	Binary Cross-entropy
output layer activation	sigmoid
hidden layer activation	tanh and relu
epochs	100
Processor	Intel Core i9
RAM	16 GB
GPU	Nvidia RTX 3080 (12 GB)

3.6.1 Proposed Time Series Prediction Model

Figure 3.4 shows the working diagram of the proposed model. It consists of 3 major blocks: data processing, training and testing the RNN or LSTM, and finally the fully connected dense layer. Data processing with feature extraction is explained in Section 3.3. After the dataset is processed and labeled, it is used to train one of the time series models. The observations in time series data are not independent. Because traditional neural networks are incapable of retaining historical information or recollection of past occurrences, they handle each observation independently. To address this problem, recurrent neural networks (RNNs) were created. RNNs bring memory into neural networks by including data point dependencies. This allows RNNs to remember information in context and learn from recurrent patterns. The primary mechanism by which RNNs achieve memory is a feedback loop within the cell. This feedback loop sets RNNs apart from ordinary neural networks. It permits information to circulate within a layer, allowing the network to remember and use data from earlier time steps. RNN sometimes suffers from vanishing and exploding gradient problems addressed in Long Short Term Memory (LSTM). It has higher memory power to remember the outputs of each node for a more extended period to produce the outcome for the next node efficiently.

$$a^{(t)} = b + Wh^{(t-1)} + Ux^{(t)} \tag{3.9a}$$

$$h^{(t)} = \tanh\left(a^{(t)}\right) \tag{3.9b}$$

$$o^{(t)} = c + Vh^{(t)} \tag{3.9c}$$

$$\hat{y}^{(t)} = \text{softmax}\left(o^{(t)}\right) \tag{3.9d}$$

Forward pass for the network is done using Eq. (3.9) where $x^{(t)}$ is the input, U is the weight associated with the input, $h(t)$ is the current state and W is its associated

weight c as the bias term, and $\hat{y}^{(t)}$ is the normalized output in terms of probability. Loss is calculated with Eq. (3.10) where L is the loss function, N is the total data points, y_i is the binary value 0 or 1 and p is their corresponding probability.

$$Loss, L = -\frac{1}{N}\sum_{i=1}^{N} \mathrm{y}_i \cdot \log\left(p\left(\mathrm{y}_i\right)\right) + \left(1 - \mathrm{y}_i\right) \cdot \log\left(1 - p\left(\mathrm{y}_i\right)\right) \tag{3.10}$$

Back-propagation is done using gradient descent on each parameter *i.e.* $\nabla_{ip} L$ where $ip = c, b, v, w, u$. After the time series model, there is a single dense layer that connects to the output layer. The iterative training will work on minimizing the error based on the ground truth labels. This in many cases may result in a vanishing gradient problem when there is a larger number of hidden layers. LSTM tends to solve the issue by introducing a *forget gate* that works on importance or weight.

$$i_t = \sigma\left(W_i \cdot [h_{t-1}, x_t] + b_i\right) \tag{3.11a}$$

$$\tilde{C}_t = \tanh\left(W_C \cdot [h_{t-1}, x_t] + b_C\right) \tag{3.11b}$$

$$f_t = \sigma\left(W_f \cdot [h_{t-1}, x_t] + b_f\right) \tag{3.11c}$$

$$o_t = \sigma\left(W_o\left[h_{t-1}, x_t\right] + b_o\right) \tag{3.11d}$$

$$h_t = o_t * \tanh\left(C_t\right) \tag{3.11e}$$

Eq. (3.11) shows the forward passing of LSTM networks, and Eq. (3.11c) is the weighted forget gate output activated by the sigmoid function. The backpropagation and loss calculation for LSTM are done similarly to RNN. We implemented the network training with Python 3.10 and TensorFlow v2. The hyperparameters used for the training are shown in Table 3.12. These parameters were chosen after trial and error for the best testing accuracy.

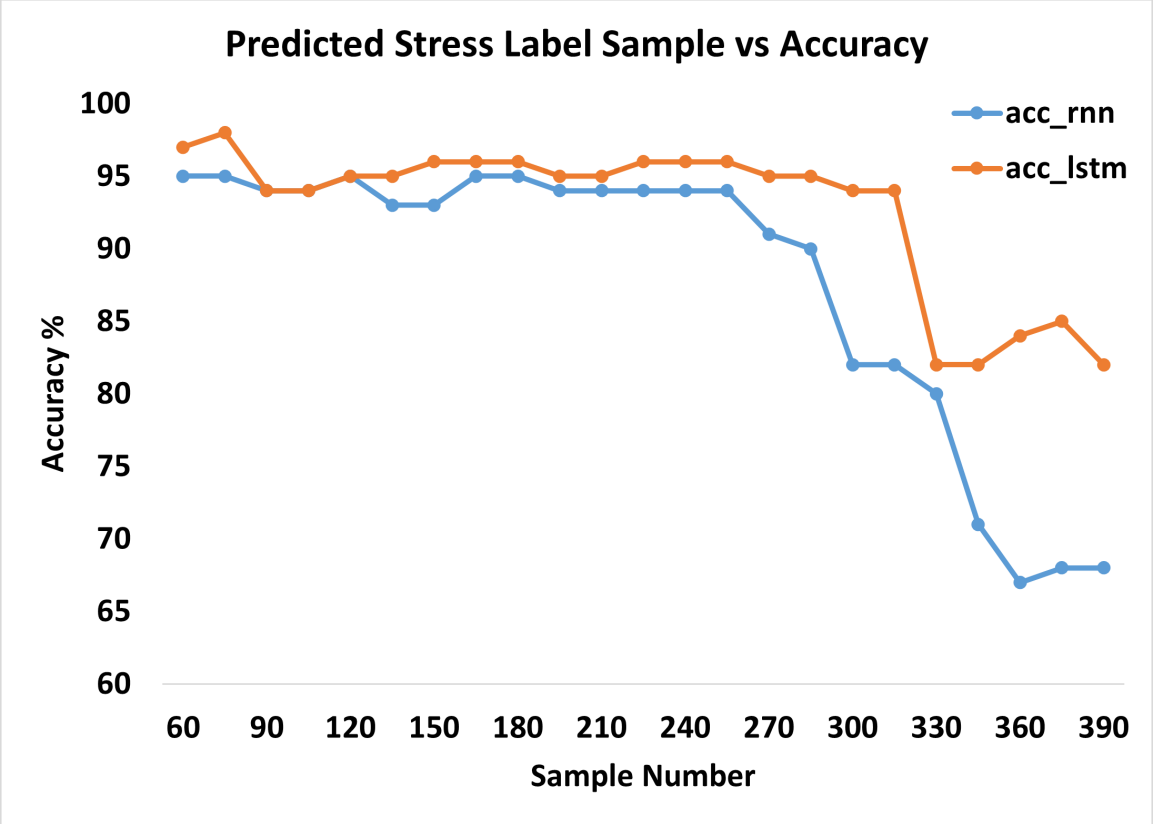


Figure 3.5: Predicted Stress Label Sample vs Accuracy

Table 3.13: Performance Metrics of Stress Prediction with LSTM and RNN for the 270th Sample (4.5 min in the future).

Method	Class	Accuracy	Precision	Recall	F1
RNN	**No-Stress**	94%	0.96	0.97	0.97
	Stress	94%	0.87	0.80	0.83
LSTM	**No-Stress**	96%	0.98	0.97	0.97
	Stress	96%	0.86	0.91	0.89

3.6.2 Result Analysis and Discussion

We evaluated the performance for both LSTM and RNN based networks with a varying length of input sequence. The experiment was done to predict the 60th window to the 390th window while the metrics were calculated and rounded up with their average. The test accuracy tends to stay stable till 4.5 minutes or the 270th sample and reduce drastically as we try to predict farther in time, as shown in Figure 3.5. Thus, we report the evaluation metrics for 4.5 minutes of prediction for both RNN and LSTM-based models in Table 3.13. We used an 80/20 split for the train and test set and randomized the fold sequence for each round of calculation, keeping the time stamp intact.

The accuracy for both models is comparable to the state of the art, and the LSTM model, as expected, showed higher recall value and F-1 score. In Table 3.14 we show a comparison of our proposed model with contemporary works for stress prediction.

It is seen that, with the improvement of ML algorithms, stress prediction tasks have also evolved. While classical machine learning models are still performing reasonably [82, 84, 88, 89] they are outperformed by task-specific time series models like LSTM and RNN [23, 85, 86]. Compared to recent works, our framework shows better evaluation metrics. This is due to the time series nature of our dataset along with the ground truth from the unbiased cortisol stress biomarker and targeted age group.

3.7 Conclusion and Future Works

The findings presented in this chapter demonstrate that digital biomarkers derived from wearable physiological signals are a viable foundation for developing context-aware stress detection and prediction models for older adults. By evaluating both feature-engineered machine learning models and CNN-based approaches, we showed that multimodal sensor features contain identifiable stress signatures that

Table 3.14: Comparison of Our Work with Contemporary Literature

Authors & Year	**Dataset**	**Stress Protocol**	**Method**	**Best Reported Metrics**
Taylor *et al.* [23] 2017	SNAPSHOT	Specific Task	Used MTL deep neural networks, Multi-task Multi-Kernel learning, Hierarchical Bayesian model 40 Participants. Best result reported with MTL-NN	Accuracy: 77.6% (MTMKL), 73.4% (HBLR), 81.5% (MTL-NN)
Reddy *et al.* [82] 2018	OSMI Survey	Daily Work	Used KNN, Logistic Regression, Random Forest, Bagging and Boosting. Best result reported with Boosting	Accuracy: 75% (Boosting), FPR: 0.25 (Boosting), Precision: 0.84 (Boosting), AUC: 0.79 (RF)
Parent *et al.* [84] 2019	Custom	Video Game	Used multi-modal data of 48 people. Best result reported with SVM	Accuracy: from 67.9% to 72.4% κ: from 0.34 to 0.42
Yu *et al.* [85] 2019	SNAPSHOT	Specific Task	Used multi-task LSTM and muilt-task CNN for 3 level stress	F-1: 0.65 (MTL-LSTM) and 0.67 (MTL-CNN)
Umematsu *et al.* [87] 2020	SNAPSHOT	Specific Task	Used Random Forest to forecast workers' daily well-being scores using physiological data	RMSE: 19.95, MAE: 13.47, CORR: 0.37(worker) RMSE: 24.76, MAE: 20.28, CORR: 0.18(Students)
Martino *et al.* [86] 2020	WESAD	TSST	Used Auto-Regressive network, Random Forest, and Least-Squares Gradient Boosting for 4 class Stress	Accuracy: from 70% to 90% for Random Forest and from 65% to 85% for LSBoost
Dai *et al.* [88] 2021	Custom	PSS and DASS	Used SVM, Random forest, gradient boosted trees, AdaBoost, and Logistic Regression on 32 participants. SVM had the best performance metrics	Accuracy: 82.6%, Precision: 0.625, Recall: 0.621, ROC: 0.79, F-1: 0.623
Kene *et al.* [89] 2021	PhysioBank	Specific Task	Used KNN, Random Forest and, support vector machine on ECG data. Best result reported with SVM	Accuracy: 83.33%, Specificity: 0.75 Recall: 0.75, F-1: 0.83
Proposed Work 2024	**Custom**	**TSST**	**Used RNN and LSTM on 40 people. Best result reported till 4.5 minutes.**	**Accuracy: 94% and 96%, Recall: 0.88 and 0.94 Precision:0.92 and 0.88, F-1: 0.90 and 0.93**

can be modeled effectively. The use of wrist-worn sensors, combined with cortisol concentration as an objective ground truth, provides a clinically relevant basis for identifying stress biomarkers and advances current methods.

In addition to detection, we investigated short-term stress forecasting using RNN and LSTM architectures to capture temporal dependencies in the physiological data. Both models achieved high predictive performance, with accuracies of approximately 94% and 96%, performing as well as or better than comparable methods in recent literature. These results highlight the potential of time-series modeling for proactive stress intervention. This is an important consideration for older adults, whose physiological regulation and stress recovery patterns may differ from younger populations.

While the results are promising, several opportunities remain for future work. Extending the models to multiclass classification may capture varying intensities of stress rather than a binary distinction, and incorporating additional contextual information may further reduce false positives and improve forecasting. This study shows that stress detection systems traditionally limited to controlled settings can be expanded to continuous, real-world monitoring in older adults. It also offers a pathway toward personalized and early intervention strategies. By grounding detection in biomarker-based, cortisol-referenced labels, the work also helps to bridge the gap between consumer-grade detection and stress assessment.

Chapter 4

Emotion Detection in Older Adults Using Physiological Signals from Wearable Sensors

This chapter investigates the problem of estimating emotion intensity using multimodal physiological features extracted from wearable sensors. While categorical emotion classification simplifies affective states into broad categories, emotion intensity estimation attempts to model subtle variations in autonomic arousal, providing a more complete understanding of emotional dynamics. The chapter evaluates multiple classical regression approaches and examines how they capture these fine-grained patterns in physiological data. This task is particularly important in affective computing applications that require precise monitoring of emotional responses in everyday environments.

The chapter also highlights the relevance of emotion monitoring for older adults, who often face increased emotional vulnerability due to loneliness, health concerns, or cognitive decline. Wearable physiological sensing offers a privacy-preserving alternative to camera-based methods that may be intrusive or difficult to use for individuals with physical or cognitive limitations.

4.1 Related Works and Motivation

Understanding emotion is vital to understanding human behavior, especially in older adults, as emotional well-being is significantly related to cognitive and physical health. Research has shown that changes in emotional patterns can serve as early indicators of neurodegenerative diseases such as Alzheimer's disease and related dementia (ADRD) [39]. They also help diagnose mental health conditions such as PTSD, depression, and anxiety disorders in veterans and other vulnerable populations. Traditional emotion recognition approaches are primarily based on facial expression analysis, speech recognition, and behavioral assessments [40, 41]. They often require camera-based monitoring or active user participation. Although effective, these methods raise concerns about privacy, intrusiveness, and real-world applicability. Recent advancements in wearable technology and physiological sensing have introduced a promising alternative to detect mental health conditions and relevant cognitive impairments [36, 42, 43]. Researchers employ both obtrusive and non-obtrusive instruments to gather physiological signals that serve as biometric indicators [44, 45, 46]. These works reveal strong correlations between physiological changes and fluctuations in mental state and highlight the versatility of physiological sensors.

Wei *et al.* [109] and Ragot *et al.* [110] employed EEG and EDA/HR, respectively, to detect broad emotional states such as positive/negative emotions and valence/arousal without relying on video data. In contrast, Hui *et al.* [27] and Kim *et al.* [111] combined physiological data with video-based methods to detect complex emotions like fear, shame, and guilt. Costa *et al.* [112] and Jiang *et al.* [113] integrated multi-modal data such as EEG, GSR, and motion data with stimuli like pictures, videos, and music to enhance emotion recognition. While these studies focus primarily on adults and young adults, Badawi *et al.* [41] addressed older adults by detecting agitation and aggression using physiological signals during indoor activities. Usually, detections are made based on biometric signals such as electrodermal activity (EDA),

heart rate variability (HRV), blood volume pulse (BVP), and skin temperature (ST). However, most existing works integrate facial emotion recognition [40, 41, 114] alongside physiological data for improved accuracy or focus on detecting a limited set of emotions [115]. Previous studies often depend on deep learning models and cloud-based processing, which significantly limit the feasibility for real-time edge deployment [116, 117]. This is of concern, particularly for older adults who may be uncomfortable or unable to participate in such assessments. Thus, the practical use of such a framework in privacy-sensitive settings such as nursing homes or veterans' care facilities is often underexplored [39]. Table 4.1 provides a detailed overview of recent research on emotion detection using physiological signals, showcasing various methodologies, detected emotions, and participant demographics.

4.2 Contributions

To address these challenges, we propose an edge-based, non-intrusive emotion detection framework that relies solely on wearable physiological sensor data without requiring cameras or facial analysis. Our proposed work focuses specifically on older adults. These groups are often overlooked in existing studies. Unlike previous research that often integrates facial recognition or video-based methods, our framework uses only physiological signals. Thus, our method eliminates privacy concerns and enables real-time edge-based deployment. This design enhances the practicality and scalability of our method for real-world use in settings such as nursing homes, clinical care, and home environments. Our contribution to this paper is as follows:

- We used a data set collected from 40 older adults, where emotions were recorded through physiological signals from the Empatica E4 and Shimmer3 GSR+ wristband.
- Participants' facial expressions were captured with live video and labeled using iMotion's Facial Expression Analysis (FEA).

Table 4.1: Overview of Recent Literature on Emotion Detection from Physiological Signals

Authors & Year	**Data Acquisition**	**Detected Emotions**	**Emotion Stimuli**	**Requires Video for Detection ?**	**Target Age Group**
Wei *et al.* [109] 2017	EEG	Positive and Negative	Pictures	✗	Adults
Hui *et al.* [27] 2018	PPG, EDA, SKT and EMG	Happiness, Anger, Fear, Disgust and Sadness	Pictures and Videos	✓	Adults
Quiroz *et al.* [118] 2018	HR	Neutral, Happy and Sad	Musics and Videos	✗	Young Adults
Heinisch *et al.* [119] 2018	EMG, EDA, TMP, EEG, BVP, RR and ECG	HPVHA, HNVLA, HPVLA HNVHA and Neutral	Pictures	✗	Young Adults
Ragot *et al.* [110] 2018	EDA and HR	Valence and Arousal	Pictures	✗	Young Adults
Xiefeng *et al.* [120] 2019	HS and ECG	Relaxed, Happy, Sad and Angry	Musics and Videos	✗	Young Adults
Costa *et al.* [112] 2019	EDA, PPG, IMU and GSR	Anger, Contempt, Disgust, Fear, Happiness, Neutral, Sadness and Surprise	Pictures	✓	Adults
Jiang *et al.* [113] 2019	EEG and Motion	Fear, Frustrated, Sad, Satisfied, Pleasant and Happy	Musics, Pictures and Videos	✓	Adults
Zhang *et al.* [26] 2020	HR, EDA, BVP, TEMP, PD, SA, and SV	Valence and Arousal	Not Mentioned	✓	Young Adults
Kim *et al.* [111] 2020	GSR and PPG	Pride, Elation, Joy, Satisfaction, Relief, Hope, Interest, Surprise, Sadness, Fear, Shame, Guilt, Envy, Disgust, Contempt and Anger	Musics, Pictures and Videos	✓	Adults
Shu *et al.* [121] 2020	HR	Neutral, Happy and Sad	Musics, Pictures and Videos	✓	Young Adults
Miranda *et al.* [122] 2021	ECG, GSR, EEG and SKT	Fear and Non-fear	Videos	✓	Adults
Chen *et al.* [123] 2021	SP	Happiness, Sadness, Anger and Fear	Videos	✗	Adults
Badawi *et al.* [41] 2024	PR, PRV, RR, SCL	Agitation and Aggression	Indoor Activities	✓	Older Adults
Proposed Work 2025	**EDA, BVP, IBI**	**Positive, Negative and Neutral**	**TSST Protocol**	✗	**Older Adults**

SA: Saccadic amplitude, SV: Saccadic velocity, HS: Heart Sound, PPG: Photoplethysmography, SKT: Skin Temperature, EMG: Electromyography, RR: Respiration Rate, PR: Pulse Rate, ECG: Electrocardiogram, EEG: Electroencephalography, EDA: Electrodermal Activity, SCL: Skin Conductance Level, IMU: Inertial Measurement Unit, PD: Pupil Dilation,HR: Heart Rate, TMP: Temperature, SC: Step Counter, BVP: Blood Volume Pressure, IBI: Inter Beat Interval, SP: Skin Potential, PRV: Pulse Rate Variability

- We detected the intensity for three distinct emotions: Neutral, Positive, and Negative with only the labeled sensor data. We used machine learning for regression tasks and reported the performance on the test set.

- By focusing on classical models rather than deep learning, we ensured computational efficiency, interpretability, and compatibility with edge-based deployment.

- We demonstrated that emotion recognition can be accurately achieved using only sensor data, eliminating the need for visual or behavioral inputs.

We have already discussed the data recording with the TSST protocol in Chapter 2 Section 2.1 Subsection 2.1.1. This chapter is organized as follows: The suggested method for emotion detection is presented in Section 4.3. The results of the study and their analysis are presented in section 4.4. Section 4.5 presents the conclusion with future direction.

4.3 Proposed Method for Emotion Detection

In this section, we will discuss our proposed method for emotion detection with machine learning-based regression models. An overview of the proposed method is shown in Figure 4.1. The three main components are (i) data preprocessing, (ii) emotion ground truth labeling with iMotion's Facial Expression Analysis (FEA), and (iii) regression with classical machine learning models. Chapter 2 Section 2.2 discusses the data preprocessing and ground truth labeling.

4.3.1 Machine Learning Models for Regression

To evaluate the performance of our framework, we implemented seven classical machine learning methods with one 10-layer Multi-Layer Perceptron (MLP) and a 3-hidden-layer Dense Neural Network. The correlation heatmap shown in Figure 4.2

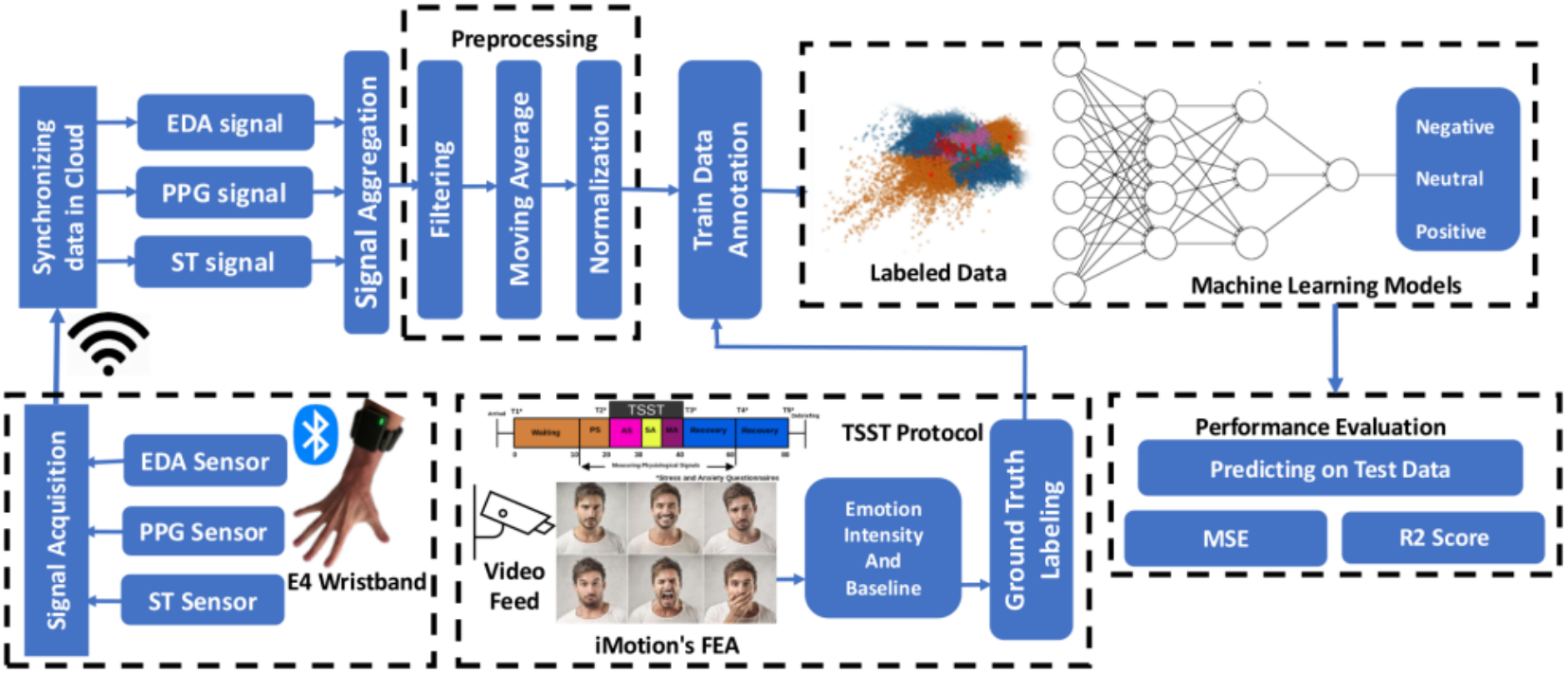


Figure 4.1: Proposed Emotion Detection Method

summarizes that traditional statistical modeling will make it difficult to make an emotion detection method. Thus, machine learning models that can handle such small correlations will prove to be well-suited for such problems.

The most simplistic regression model is the linear regression. It assumes a direct linear relationship between physiological features and emotional intensity. It estimates coefficients that best fit the data by minimizing the sum of squared errors. Ridge regression is a linear model that incorporates L2 regularization, which penalizes large coefficients to prevent overfitting. Bayesian ridge regression extends ridge regression by incorporating a probabilistic framework, which assigns prior probability distributions to the model coefficients.

K-Nearest Neighbors (KNN) Regression is a non-parametric model that estimates emotion intensity based on the average intensity values of the K most similar physiological responses. This approach assumes that individuals with similar physiological reactions experience similar emotional intensities. Decision tree regression is a rule-based model that recursively splits the data into subsets based on the feature values that minimize the prediction error at each node. Another prominent method is the random forest regression, which is an ensemble learning method. It constructs multiple decision trees during training and averages their predictions for a more accurate and robust outcome.

Gradient boosting regression is a more complex model. It is an advanced ensemble technique that sequentially builds multiple weak learners (typically decision trees), where each new tree corrects the errors of the previous ones. This method reduces both bias and variance, leading to highly accurate predictions. Multi-Layer Perceptron (MLP) Regression is a type of feedforward neural network consisting of multiple hidden layers. Each layer processes the input data using weighted connections and nonlinear activation functions. Finally, a Dense Neural Network (DNN) consists of multiple fully connected layers that transform input features through non-linear activation functions. By learning hierarchical representations,

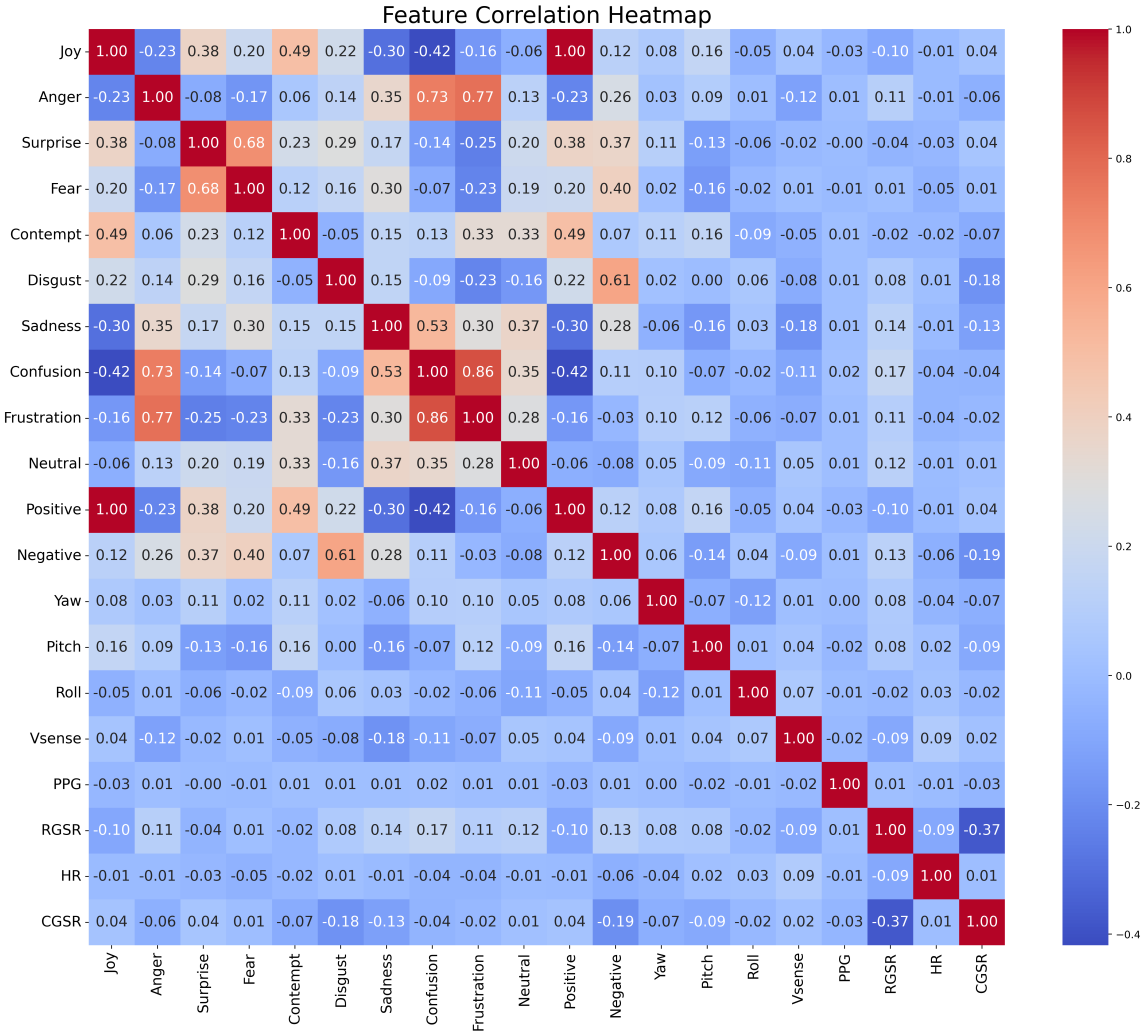


Figure 4.2: Correlation Heatmap of the Features and Emotion of Dataset

DNNs can capture complex interactions between physiological signals and emotional intensity levels.

4.3.2 Model Training and testing

As we had participant-specific data, there were multiple ways to generalize the machine learning models based on the dataset. To avoid data leakage, we split the dataset based on participants before normalizing them. We performed *Leave One Sample Out (LOSO)* cross-validation, keeping one subject for testing and others for training. We used some widely used regression models as our base machine learning model mentioned in Section 4.3.1. For each of the regression models, the set of hyperparameters used is shown in Table 4.2. The values for these parameters were chosen based on the best-performing set on test data after multiple trials and errors.

4.4 Result Analysis and Discussion

In this section, we will evaluate the performance of machine learning models with our dataset with respect to the estimated ground truth mentioned in Section 2.2.3. We have evaluated the model's performance using r^2 score and MSE with the equations shown in Eq. (4.1) and (4.2), where y_i = actual value, $\hat{y}_i$= predicted value, and $\bar{y}_i$ = mean of actual value. The r^2 score of a model says how much of the variability in the target that model can explain in terms of normalized distance. Again, MSE refers to the squared difference between the model's predicted values and the actual values. The scores are expressed as the mean of all folds for LOSO cross-validation.

$$R^2 \; score = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2} \tag{4.1}$$

$$\text{MSE} = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 \tag{4.2}$$

Table 4.2: Hyper-parameters Used to Train Machine Learning Models

Machine Learning Model	**Hyper-parameters**
Random Forest	no of estimators=100, criterion=squared error
Dense Neural Network	no of hidden layers=3, loss=mse, output layer activation=ReLU, optimizer=Adam, epochs=50, Learning rate = 0.0001
Decision Tree	criterion=squared error
K-Nearest Neighbors (KNN)	no of neighbors=3, weights=uniform, distance metrics=minkowski
Gradient Boosting	no of estimators=100, Learning rate=0.1, loss=squared error,criterion=friedman mse
Multi Layer Perceptron	no of hidden layers=10, loss=mse, output layer activation=ReLU, optimizer=Adam, Learning rate = 0.001, maximum iteration=500
Ridge Regression	maximum iteration=1000, tolerance=10^{-4}, solver=svd
Bayesian Ridge Regression	maximum iteration=1000, tolerance=10^{-3}, $\alpha_1, \alpha_2, \lambda_1, \lambda_2 = 10^{-6}$
Linear Regression	N/A

The results indicate that Random Forest outperforms other models, as evidenced by its highest r^2 scores in Table 4.3 and lowest Mean Squared Error (MSE) in Table 4.4 across all emotion categories. That is because the ensemble nature of Random Forest helps reduce overfitting and enhances generalization. This indicates that the Random Forest model was able to explain a substantial portion of the variance in emotional states with reasonable confidence.

In contrast, simpler models such as Ridge Regression, Bayesian Ridge Regression, and Linear Regression perform poorly. They showed significantly lower r^2 scores with high MSE. This implies that linear models fail to capture the intricate dependencies between physiological signals and emotion states. Since physiological signals are often highly nonlinear and affected by multiple interacting factors, simpler models struggle to provide meaningful predictions.

Neural network-based models, such as the Dense Network and Multi-Layer Perceptron (MLP), perform moderately well. While they achieve better r^2 scores than linear models, their higher MSE values suggest a higher degree of variance in their predictions. However, their ability to outperform purely linear models suggests that with further optimization and larger datasets, deep learning approaches could become more competitive in emotion detection from physiological signals with more complexity, as that would make them harder to deploy on the edge.

4.5 Conclusion

This chapter demonstrates the feasibility of detecting emotional intensity using physiological signals alone. This framework eliminates the need for camera-based or facial-recognition approaches that may be intrusive or impractical for many users. The results show that classical machine learning models can effectively predict emotional states, highlighting the value of wearable sensing as a privacy-preserving and edge-compatible alternative. These findings are particularly relevant for populations such as individuals with Alzheimer's Disease and Related Dementias (ADRD),

Table 4.3: r^2 Scores for Various ML Models (Higher is better)

ML Model	**Negative**	**Neutral**	**Positive**
Random Forest	0.782	0.7636	0.8033
Dense Network	0.5657	0.5339	0.5813
Decision Tree	0.5377	0.5253	0.5762
K-Nearest Neighbors (KNN)	0.5217	0.5021	0.5673
Gradient Boosting	0.3545	0.3644	0.2997
Multi Layer Perceptron	0.3088	0.2568	0.2655
Ridge Regression	0.0752	0.0371	0.0527
Bayesian Ridge Regression	0.0752	0.0370	0.0526
Linear Regression	0.0752	0.0370	0.0526

Table 4.4: MSE for Various ML Models (Lower is better)

ML Model	**Negative**	**Neutral**	**Positive**
Random Forest	0.0006	0.0014	0.0019
Dense Network	0.0012	0.0029	0.0030
Decision Tree	0.0012	0.0030	0.0030
K-Nearest Neighbors (KNN)	0.0013	0.0031	0.0044
Gradient Boosting	0.0017	0.0042	0.0050
Multi Layer Perceptron	0.0018	0.0047	0.0052
Ridge Regression	0.0025	0.0060	0.0067
Bayesian Ridge Regression	0.0025	0.0060	0.0067
Linear Regression	0.0025	0.0060	0.0067

veterans with Post-Traumatic Stress Disorder (PTSD), and others who experience cognitive or communicative limitations. By relying solely on passive physiological measurements, this approach allows continuous monitoring of emotional well-being in clinical, residential, and assisted-living environments without requiring active user engagement.

The implications of these results extend to both healthcare and human-centered AI, providing a foundation for scalable emotion-aware systems that can support early detection, intervention, and personalized care. While the current models show strong predictive capability, future work may benefit from exploring deep learning architectures, transfer learning techniques, and larger, more diverse datasets to further improve predictive ability. Overall, this study establishes a reliable baseline for real-time, privacy-conscious emotion detection using wearable physiological data and serves as a starting point toward advanced emotion-tracking technologies in health-focused applications.

Chapter 5

Quantum Hybrid Support Vector Machines for Stress and Emotion Detection in Older Adults

This chapter introduces a quantum-hybrid machine learning framework designed to enhance stress and emotion detection using physiological signals. These models capture complex correlations by embedding data into high-dimensional quantum feature spaces through parameterized quantum circuits. In most cases these may be difficult or inefficient for classical kernels to model. The chapter describes the construction of quantum feature maps, the role of entanglement in generating expressive representations, and the use of fidelity kernels to quantify similarity between encoded states. These ideas form the basis for applying quantum computing tools to physiological signal analysis.

A secondary focus of the chapter is the application quantum representations may offer for older adult populations, whose physiological responses often contain higher variability and nonlinear dynamics. In such cases, richer feature embeddings may reduce misclassification and improve robustness to noise. This chapter establishes

the theoretical and practical considerations for using quantum kernels in hybrid architectures that work in the boundary of modern time stress and emotion detection.

5.1 Related Works and Motivation

Excessive stress can harm a person's mental health, leading to various medical conditions. It can raise the risk of mental and physical disorders, such as immune system issues, cardiovascular and metabolic system impairment, and so on. Besides, stress is connected to cognitive decline, which over time could increase the onset of Alzheimer's disease in older adults, as well as physiological irregularities. Multiple studies suggest that experiencing a longer period of stressful situations can contribute to the deterioration of mental and physical health, as well as mortality [124, 13]. Thus, early detection of stress is necessary for successful intervention. Recent advances in machine learning and quantum computing (QC) have created new opportunities to solve such problems with artificial intelligence (AI) [60, 125]. Quantum machine learning (QML) is a promising field generating active research across multiple domains. Researchers have increasingly used various quantum algorithms for real-life healthcare applications. In this domain, quantum machine learning has grown as an emerging AI framework. It attracts interest from both academic and medical communities. Researchers made significant experimental advances in classifying and diagnosing stress. There have been recent advancements made in the development of quantum support vector machines (QSVM), quantum neural networks, variational quantum classifiers (VQC), error reduction tactics, and advanced pre-processing methods [60, 126, 127, 128, 129]. QML uses the encoded classical dataset, which is transformed into quantum states. There are several benefits of encoding data into quantum states. It helps perform quantum computation with classical data. For linear classifiers, kernel-based methods have also been effective in providing dataset uniformity. Similar to classical machine learning, researchers are currently working to create reliable quantum algorithms that are suited to classifying problems unique to

mental health, such as stress detection [130, 131]. Although, there have been multiple experiments on classical algorithms that can detect and predict stress [38, 44, 39, 132]. In that regard, hybrid QC and ML algorithms have been applied to multiple datasets [133, 134, 135, 136]. These algorithms combine classical algorithms with the benefits of quantum computing approaches. Again, the application of such a framework for older adults is under-explored [38, 44, 137] while there are a lot of works focusing on young adults [84, 23]. Thus, developing innovative machine learning approaches that take advantage of QC's specific features for healthcare applications is a timely necessity. Their performances also need to be evaluated for real-life scenarios that can justify the viability of such a framework for older adults.

On the other hand, emotional well-being plays a critical role in the overall health of older adults. Recent works show substantial evidence indicating strong associations of emotional states with both cognitive and physical functioning. Deviations in emotional patterns have been identified as early markers of neurodegenerative conditions such as Alzheimer's disease and related dementias (ADRD) [39]. Accurate emotion detection is thus essential for the diagnosis and monitoring of psychiatric conditions, including post-traumatic stress disorder (PTSD), depression, and anxiety—particularly among older adults and other vulnerable populations.

Traditional emotion recognition methodologies have predominantly relied on facial expression analysis, speech processing, and behavioral assessments. While these approaches have demonstrated effectiveness, they often necessitate camera-based systems or active user engagement, posing challenges related to user privacy, intrusiveness, and scalability in naturalistic settings [40, 41]. Recent advances in wearable technologies and physiological sensing offer a compelling alternative for emotion recognition through continuous, noninvasive data collection. Physiological signals—such as electrodermal activity (EDA), heart rate variability (HRV), blood volume pulse (BVP), and skin temperature—serve as robust indicators of emotional and cognitive states. Prior research utilizing wearable devices like the Empatica E4 has demonstrated promising results in real-time detection of stress, anxiety, and

cognitive load during both clinical and educational tasks [36, 42, 43]. In these tasks, both obtrusive and non-obtrusive instruments were used to gather physiological signals that serve as biometric indicators [44, 45, 46].

Despite these advancements, the high dimensionality, temporal complexity, and noise inherent in physiological time-series data present substantial challenges for classical machine learning models, which are often computationally intensive and limited in their capacity to generalize across heterogeneous populations. In response to these limitations, Quantum Machine Learning (QML) has emerged as a novel computational paradigm capable of addressing the complexity of physiological data with enhanced efficiency and potentially greater expressiveness. By harnessing principles of quantum mechanics, QML models offer the potential to process high-dimensional data with reduced resources and improved representations.

Table 5.1 provides some of the recent works on emotion detection in both the classical and quantum domains. It showcases various methodologies, detected emotions, and participant demographics. While these studies focus primarily on adults and young adults, in order to address older adults, Badawi *et al.* [41] used physiological markers to identify agitation and aggression during indoor activities. Biometric signals like skin temperature (ST), blood volume pulse (BVP), heart rate variability (HRV), and electrodermal activity (EDA) are typically used to make detections. However, for increased accuracy, the majority of current works combine facial emotion recognition [40, 41, 114] with physiological data, or they concentrate on identifying a small number of emotions rather than a wide variety of affective states [115]. Prior research frequently relies on cloud-based processing and deep learning models, which severely restricts the viability of real-time edge deployment [116, 117]. This is concerning, especially for older persons who may feel uncomfortable or unwilling to participate in such assessments. The practical application of such a framework in privacy-sensitive settings, such as nursing homes or veterans' care facilities, is typically underexplored [39].

Table 5.1: Recent Literature on Emotion Detection

Method	Authors & Year	Data Acquisition	Detected Emotions	Emotion Stimuli	Requires Video for Detection ?	Target Age Group of Participants
Classical	Zhang *et al.* [26] 2020	HR, EDA, BVP, TEMP, PD, SA, and SV	Valence and Arousal	Not Mentioned	✓	Young Adults
	Shu *et al.* [121] 2020	HR	Neutral, Happy and Sad	Musics, Pictures and Videos	✓	Young Adults
	Miranda *et al.* [122] 2021	ECG, GSR, EEG and SKT	Fear and Non-fear	Videos	✓	Adults
	Chen *et al.* [123] 2021	SP	Happiness, Sadness, Anger and Fear	Videos	✗	Adults
	Badawi *et al.* [41] 2024	PR, PRV, RR, SCL	Agitation and Aggression	Indoor Activities	✓	Older Adults
Quantum	Li *et al.* [145] 2021	Texts, Audio and Videos	Happy, Sad, Neutral, Angry, Excited, Frustrated	Acts from TV	✓	Young Adults
	Singh *et al.* [139] 2022	Videos	Anger, Disgust, Fear, Joy, Sadness, and Surprise	Posed	✓	Young Adults
	Li *et al.* [146] 2023	Texts, Audio and Videos	Neutral, Happy and Sad	Speech and Interaction	✓	Young Adults
	Mai *et al.* [138] 2024	EEG	Anger, Disgust, Fear, Happiness, Sadness and Surprise	Music Videos	✗	Young Adults
	Liu *et al.* [147] 2024	Texts, Audio and Videos	Happy, Sad, Neutral, Angry, Excited, Surprise	Acts from TV	✓	Adults
	Golchha *et al.* [141] 2025	Texts, Audio and Videos	Anger, Disgust, Happiness, Fear, Sadness, Surprise and Neutral	Posed	✗	Not Specific
	Singh *et al.* [142] 2025	Texts, Audio and Videos	Anger, Disgust, Fear, Happiness, Sadness and Surprise	Indoor Activities	✓	Not Specific
	Proposed Work 2025	**EDA, BVP, IBI**	**Positive, Negative and Neutral**	**TSST Protocol**	✗	**Older Adults**

SA: Saccadic amplitude, SV: Saccadic velocity, SKT: Skin Temperature, RR: Respiration Rate, PR: Pulse Rate, ECG: Electrocardiogram, EEG: Electroencephalography, EDA: Electrodermal Activity, SCL: Skin Conductance Level, PD: Pupil Dilation, HR: Heart Rate, TMP: Temperature,SC: Step Counter, BVP: Blood Volume Pressure, IBI: Inter Beat Interval, SP: Skin Potential, PRV: Pulse Rate Variability

In parallel, researchers have moved toward using actual quantum circuits for emotion recognition tasks [138, 139, 140]. Their studies employ Variational Quantum Circuits (VQCs) to process data from various sources such as EEG signals and facial expressions. These hybrid quantum-classical models are designed to extract meaningful features and perform classification that applies quantum computing principles to wearable and real-time emotional monitoring systems. Additionally, researchers further broaden the scope of QML in emotion detection [141, 142, 143, 144] using quantum-inspired neural networks. They integrated quantum-enhanced processing with classical convolutional neural networks for facial emotion recognition. These efforts reflect a collective move towards harnessing the strengths of both quantum and classical paradigms to enhance emotion detection capabilities across various data types and application scenarios.

5.2 Contributions

To address the issue, we experiment with the kernel-based Quantum Hybrid Support Vector Machine (QHSVM) on a real-life dataset of older adults in a constrained environment. Our contributions in this work are:

- We used a dataset of 40 older adults whose emotions were recorded using facial expression and physiological signals from the Empatica E4 and Shimmer3 GSR+ wristbands.
- Live video was used to capture the participants' facial expressions, which were then labeled using iMotion's Facial Expression Analysis (FEA) module.
- Our proposed hybrid framework looks at the stress detection problem as an anomaly detection problem.
- The inner product kernel finds the fidelity between features to capture the inherent pattern that distinguishes the anomaly from the baseline.

- Using solely labeled sensor data, we detected the intensity of three distinct emotions: neutral, positive, and negative. We showed that emotion detection can be accurately performed using simply sensor data, without the need for graphical or cognitive inputs.

- We performed classification tasks using both classical machine learning and hybrid quantum machine learning, and we provided our findings on the test set for different numbers of samples.

- Finally, we show a performance analysis with experimental results on quantum simulation.

- We also showed that incorporating quantum machine learning can improve the overall performance, reducing the false positive detection.

The chapter is mainly divided into two parts as follows: Section 5.3 describes the proposed method for stress detection with QML. The proposed method for emotion detection is provided in Section 5.4. Section 5.4.3 summarizes the study's findings and analyses. Section 5.5 presents the conclusion and future prospects.

5.3 Proposed Method for Stress Detection

The procedure for this hybrid SVM method consists of the following stages: First, we perform feature selection and preprocessing in the classical machine. Then, we compute the fidelity between features with the inner product of the probability taken from the quantum kernel circuit. Finally, use this information to train and classify a OneClass SVM in the classical machine. The whole process is shown in Figure 5.1.

5.3.1 Data Collection, Preprocessing and Feature Selection

The physiological data collection with the TSST protocol is explained in detail in 2 Section 2.1.2. The preprocessed data then goes to the feature extraction states

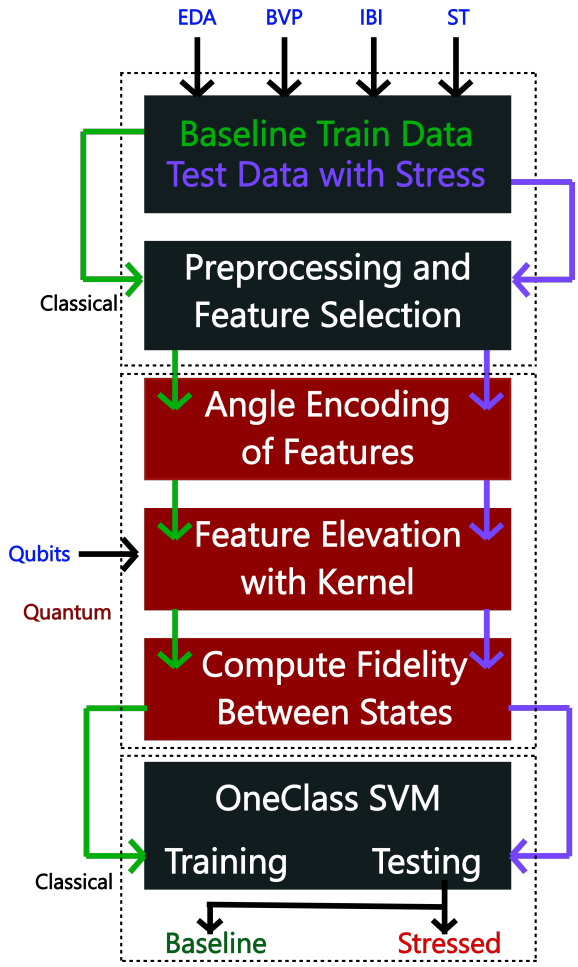


Figure 5.1: Pipeline of the Hybrid-Quantum SVM Anomaly Detection

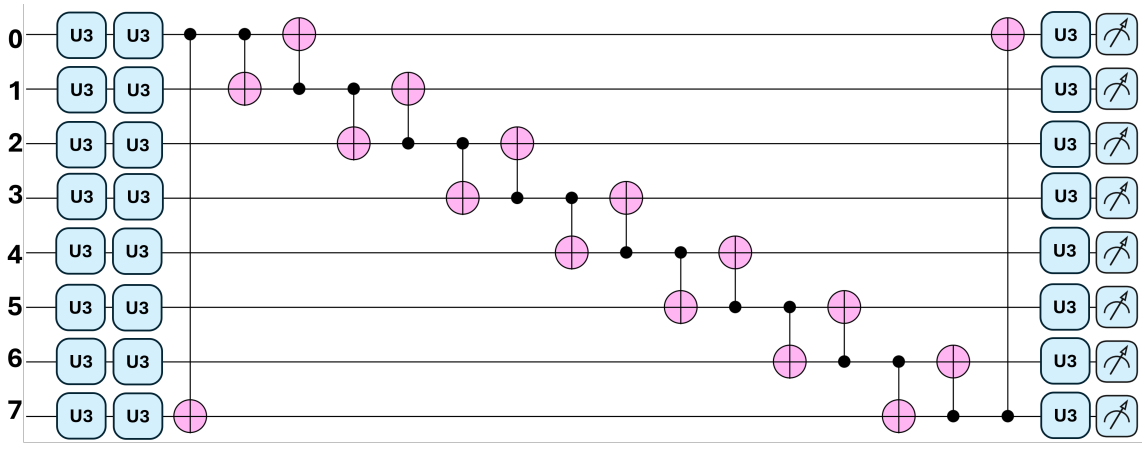


Figure 5.2: Quantum Circuit for Feature Mapping with 8 qubits

mentioned in 3 Section 3.4.1. But encoding 60 features to quantum states to perform the kernel computation is a very computationally expensive task, and the runtime will also be impractical. For each set of additional features, the complexity increases quadratically. One way to address that would be parallel computation, but the increase in features also increases the circuit width, keeping the circuit depth the same. This prevents the parallel computation of the kernel. Again, some less correlated features tend to throw the machine learning models off from convergence, resulting in poor accuracy. Thus, to accommodate the highly correlated features with reasonable complexity, we select the best 8 and 12 features with the feature selection technique. To select the best features, a decision tree is trained using all training data. The tree can separate each sample of the features based on the splitting criteria or ground truth. Later the features are scored in accordance with their ability to reduce the error or impurity. For decision trees or random forest-based models, the error is calculated with the *Gini Index* and *Entropy*. For our work, we ranked the features according to their relative decrease in Gini impurity, as shown in Eq. (5.1), (5.2). In this equation, I represents the impurity estimated for the probability P of a sample occurrence in node N, while I_0 represents the impurity calculated from the preceding node N_0.

$$I(n) = 1 - \sum P^2 \tag{5.1}$$

$$\Delta I = I_0 - \left(\frac{N_{left}}{N_0} \times I_{left} + \frac{N_{right}}{N_0} \times I_{right} \right) \tag{5.2}$$

Again, before feature extraction and selection, we randomly sample 2000 baseline data points for training and 1500 data points with a combination of baseline and stress labels for testing.

5.3.2 Feature Mapping with Fidelity Quantum Kernel

The quantum kernel circuit is made with three steps. The first part is the encoding, where each line is initialized as a 0 state of qubits. Here, the classical data is encoded in a way that each qubit can represent two features. This encoding scheme named Dense Angle Encoding was proposed by [55], where N features can be encoded into $N/2$ qubits. Figure 5.2 shows the kernel circuit. $U3$ is a unitary rotation consisting of a single qubit rotation that works as angle encoding. The R_x gate applies a rotation about the x-axis of the Bloch sphere, mathematically defined as $R_x(\theta) = e^{-i\theta X/2}$, where X is the Pauli-X matrix and θ is the angle of rotation, which can be set as proportional to x_1. Similarly, the R_z gate rotates the qubit around the z-axis and is defined as $R_z(\phi) = e^{-i\phi Z/2}$, with Z being the Pauli-Z matrix and ϕ proportional to x_2. By applying $R_x(x_1)$ followed by $R_z(x_2)$ to each qubit, the two features x_1 and x_2 are mapped to the qubit's quantum state. Afterward, the series of $CNOT$ gates entangles a feature with the rest of them. After the measurement, the probability of keeping initial states is fed into an inner-product kernel based on quantum fidelity. It is expected to measure a low fidelity and higher for the baseline. The fidelity of the same feature was normalized to 1. The Fidelity kernel captures the high-dimensional similarities among the features. This circuit is constructed following the work of [148].

5.3.3 Training and Testing with the One-Class Support Vector Classifier

In the training phase, the OneClass SVM finds the ideal hyperplane between classes by identifying the support vectors. In kernel-based OneClass SVM, similar data points are bounded by a decision hyperplane where the anomalies are kept out. If $\{x_r\}_{r=1}^{N}$ represents the data from a single class, the decision function for a one-class SVM can be represented as Eq. (5.3).

$$f(x) = sign\left(\sum_{r=1,s=1}^{N} \alpha_r \times k(x_r, x_s) + b - \rho\right) \tag{5.3}$$

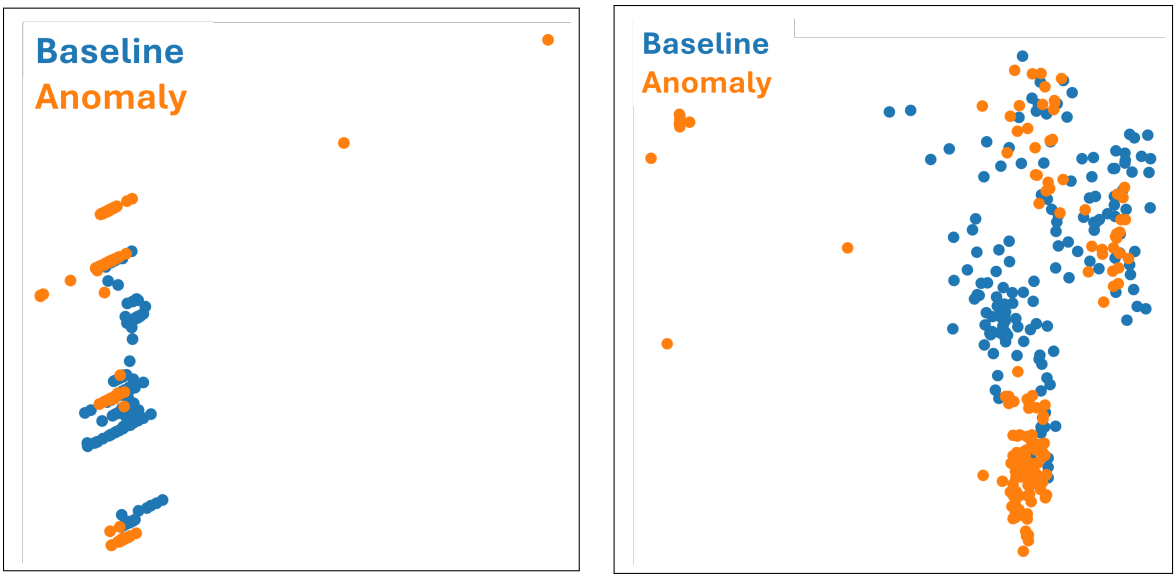
Here, $k(x_r, x_s)$ is the kernel function, and x_r and x_s are the input vectors. α_r are the tunable constants. b is the bias term and ρ is the threshold. The inner product kernel shape for training samples is $N_{train} \times N_{train}$ with a symmetry by the diagonal, and the shape for the test kernel is $N_{test} \times N_{train}$. Projecting the testing kernel in a 2D space can help visualize the difference between the classical and quantum kernels.

It is seen from Figure 5.3 that, after the kernel computation, both the classical and quantum kernels grouped the test samples. For classical kernel 5.3a there are a lot of overlaps that make drawing the decision boundary harder for any classifier. On the other hand, with high-dimensional correlation, the quantum kernel 5.3b has fewer overlaps. The 2D projection thus helped us hypothesize that the quantum SVM can expect better performance.

5.3.4 Result Analysis

Our design for the quantum circuit and the simulation was performed using PennyLane [149] with *qiskit aer* backend [150] with no noise model. In this section, we describe the experimental result achieved with our stress dataset. We also compare the performance with the equivalent classical model having a similar number of features. To properly identify the benefits of using a fidelity-based quantum kernel, a table of various performance metrics is provided in Table 5.2.

The table shows proposed QSVM performed better than classical SVM in all trials. It works best with 12 features and 6 qubits with 79% accuracy and an F-1 score of 0.78. Again, another observation from the experiment is the low precision and high recall value for quantum SVM. This is particularly useful for medical diagnosis and anomaly detection, where a higher recall value provides an advantage. Thus, the proposed QSVM model proved to be more viable for stress detection. In such a context missing a positive case can be damaging, and a higher recall value prevents



(a) Classical Kernel

(b) Quantum Kernel

Figure 5.3: 2D Projection of Test Kernel Matrix with Classical and Quantum Kernel. (*blue* represents Baseline or Normal class and *orange* represents Anomaly or Stress class)

Table 5.2: Performance Metrics of Quantum SVM Model and Classical SVM Based on Feature (and qubit) counts for Stress Detection.

Method	**Features**	**Accuracy**	**Precision**	**Recall**	**F1**
Quantum	**8 (4-Qubits)**	Max: 70% Avg: 61%	Max: 65% Avg: 63%	Max: 72% Avg: 64%	Max: 0.75 Avg: 0.64
	12 (6-Qubits)	Max: 79% Avg: 64%	Max: 80% Avg: 64%	Max: 79% Avg: 64%	Max: 0.78 Avg: 0.63
Classical	**8**	Max: 64% Avg: 57%	Max: 76% Avg: 59%	Max: 65% Avg: 59%	Max: 0.66 Avg: 0.59
	12	Max: 77% Avg: 62%	Max: 77% Avg: 64%	Max: 75% Avg: 65%	Max: 0.75 Avg: 0.65

that. Comparing this to a classical counterpart, the low recall value put the model at a disadvantage.

5.4 Proposed Method for Emotion Detection

In this section, we will discuss our proposed method for emotion detection with machine learning-based regression models. An overview of the proposed method is shown in Figure 5.4. The three main components are (i) data preprocessing, (ii) emotion ground truth labeling with iMotion's Facial Expression Analysis (FEA), and (iii) multiclass classification with classical and hybrid quantum machine learning models.

5.4.1 Data Preprocessing and Ground Truth Labeling

Emotion ground truth labeling and FEA modules are discussed in 2 Section 2.2 Subsection 2.2.3 and Subsection 2.2.2 in detail. This section will highlight the model training and testing with performance analysis.

5.4.2 Model Training and Testing

To assess the performance of our framework, we used seven traditional machine learning approaches to compare against the proposed hybrid quantum machine learning model. They are Random Forest, K-Nearest Neighbors, Decision Tree, Gradient Boosting, Naive Bayes, Support Vector Machine, and Logistic Regression. As a base hybrid model, we use support vector machine with a quantum kernel. As we had participant-specific data, there were multiple ways to generalize the machine learning models based on the dataset. To avoid data leakage, we split the dataset based on participants before normalizing it. We performed 5-fold cross-validation, keeping 20% for testing and 80% for training. For each of the models, the set of

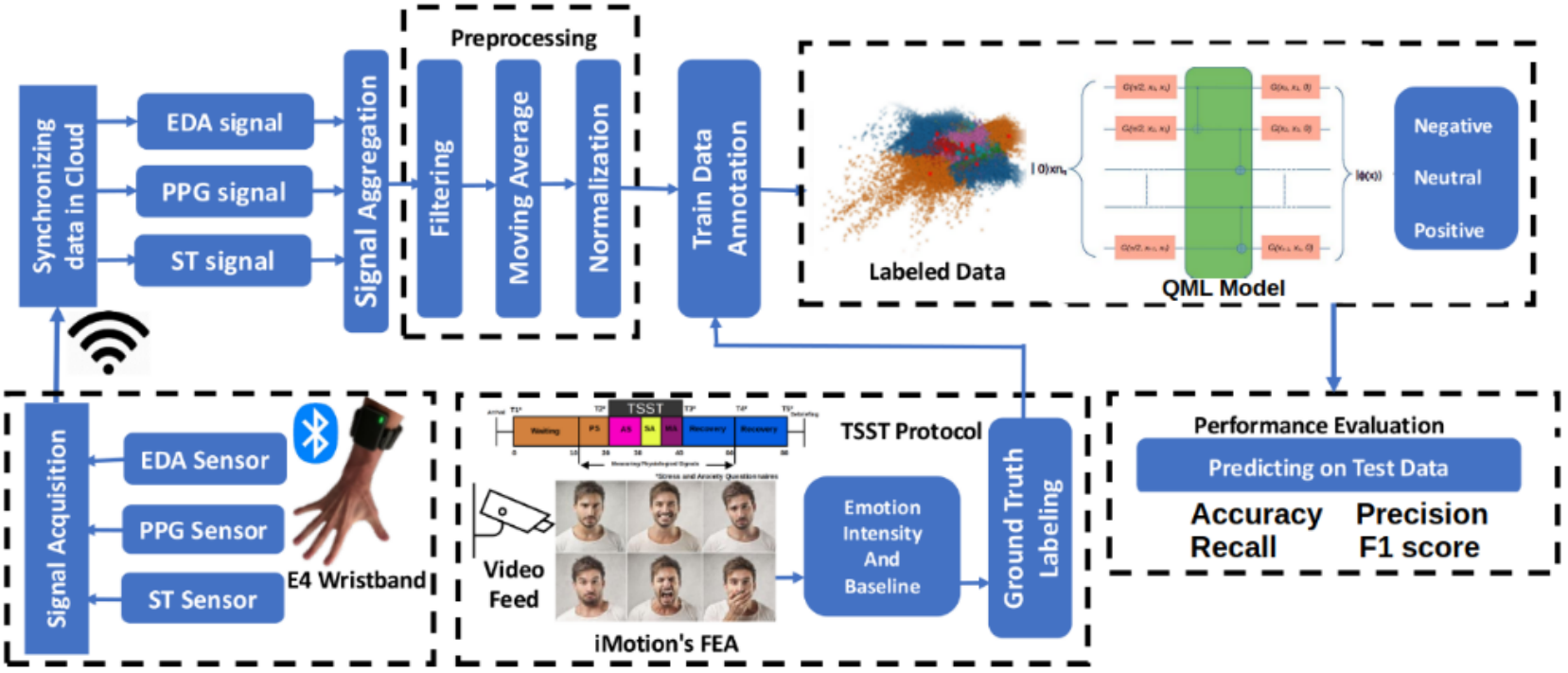


Figure 5.4: Proposed Emotion Detection Method with QML

Table 5.3: Hyper-parameters Used to Train Machine Learning Models

Machine Learning Model	**Hyper-parameters**
Random Forest	no of estimators=100, criterion=squared error
Decision Tree	criterion=squared error
K-Nearest Neighbors (KNN)	no of neighbors=3, weights=uniform, distance metrics=minkowski
Gradient Boosting	no of estimators=100, Learning rate=0.1, loss=log_loss, max depth=3
Multi Layer Perceptron	no of hidden layers=10, loss=mse, output layer activation=ReLU, optimizer=Adam, Learning rate=0.001, maximum iteration=500
Support Vector Machine	Kernel=rbf, tolerance=10^{-4}
Logistic Regression	solver=lbfgs
Quantum SVM	**Kernel=quantum_kernel, tolerance=10^{-4}, n_qubits=4**

hyperparameters used is shown in Table 5.3. The values for these parameters were chosen based on the best-performing set on test data after multiple trials and errors.

Next, the quantum hybrid support vector machine was trained with the kernel calculated from the training and test sets. Given the quantum kernel matrix K, an SVM is used to perform the final classification. The optimization problem is formed as Eq. (5.4) when $0 \leq \alpha_i \leq C, \quad \sum_{i=1}^{N} \alpha_i y_i = 0$.

$$\max_{alpha} \sum_{i=1}^{N} \alpha_i - \frac{1}{2} \sum_{i,j=1}^{N} \alpha_i \alpha_j y_i y_j k(x_i, x_j) \tag{5.4}$$

where α_i are the coefficients, $y_i \in \{-1, 1\}$ are the class labels, and C) is a regularization parameter. During inference, the decision function for a new input x is given by $f(x) = \text{sign}\left(\sum_{i=1}^{N} \alpha_i y_i k(x_i, x) + b\right)$, where b is the bias term.

We utilized *Python 3.10* and the *Qiskit* library to compute quantum kernel matrices for emotion detection. The quantum kernel is constructed using the *FidelityQuantumKernel* from *Qiskit*'s *machine_learning* module that computes the inner products of these quantum states. The simulation is executed on a classical backend with *AerSimulator* using the *statevector* method to compute the fidelity between states. The resulting training and test kernel matrices are forwarded to SVM and optimized with the *SVC* package of the *scikit-learn* module.

5.4.3 Result Analysis and Discussion

In this section, we will evaluate the performance of machine learning models with our dataset with respect to the estimated ground truth mentioned in Section 2.2.3.

The results presented in Table 5.4 provide a comparative analysis of traditional machine learning algorithms for the emotion classification task with 2,000 samples. Random Forest stands out as the top-performing model by consistently achieving the highest F1 scores across the Negative (0.685), Neutral (0.851), and Positive (0.761) classes. These results highlight its strong balance between precision and recall. Tree-based approaches, such as Random Forest and Decision Tree, demonstrate superior

Table 5.4: Performance of Classical ML Models for 2000 Samples (avg of 5 folds)

ML Algorithm	**Negative**				**Neutral**				**Positive**			
	Accuracy	**Precision**	**Recall**	**F1 Score**	**Accuracy**	**Precision**	**Recall**	**F1 Score**	**Accuracy**	**Precision**	**Recall**	**F1 Score**
Random Forest	0.796	0.768	0.619	0.685	0.901	0.925	0.788	0.851	0.885	0.853	0.688	0.761
K-Nearest Neighbors	0.762	0.662	0.600	0.629	0.863	0.750	0.680	0.713	0.846	0.735	0.667	0.699
Decision Tree	0.747	0.614	0.615	0.615	0.846	0.696	0.697	0.697	0.830	0.682	0.684	0.683
Gradient Boosting	0.699	0.710	0.203	0.316	0.792	0.804	0.230	0.358	0.776	0.789	0.226	0.351
Naive Bayes	0.655	0.434	0.201	0.274	0.742	0.492	0.227	0.311	0.728	0.482	0.223	0.305
Support Vector Machine	0.685	0.720	0.132	0.223	0.777	0.816	0.149	0.253	0.761	0.800	0.146	0.248
Logistic Regression	0.664	0.635	0.036	0.068	0.753	0.720	0.041	0.077	0.738	0.705	0.040	0.075

Table 5.5: Performance of Hybrid Quantum SVM Model with Belis Kernel (avg of 5 folds)

No of Samples	**Negative**				**Neutral**				**Positive**			
	Accuracy	**Precision**	**Recall**	**F1 Score**	**Accuracy**	**Precision**	**Recall**	**F1 Score**	**Accuracy**	**Precision**	**Recall**	**F1 Score**
train: 640; test: 160	0.800	0.853	0.860	0.851	0.863	0.751	0.803	0.775	0.892	0.878	0.938	0.905
train: 1600; test: 400	0.815	0.887	0.864	0.872	0.874	0.790	0.864	0.824	0.881	0.896	0.898	0.897

performance over linear models, particularly for the Neutral and Positive classes. This suggests that tree-based models work better to capture the non-linear relationships inherent in emotion data. In contrast, logistic regression and support vector machine (SVM) with RBF kernel exhibit relatively high precision but notably low recall. This imbalance leads to reduced F-1 scores.

On the other hand, Table 5.5 shows that hybrid quantum SVMs utilizing the Belis kernel exhibit clear performance advantages. The quantum model surpasses classical models in most classes, even with smaller samples. With only 640 training samples, the quantum SVM achieves an F1 score of 0.905 for the positive class that exceeds the best classical result of 0.761 from Random Forest. When the training set increases to 1,600 samples, the hybrid quantum model's robustness and generalization further improve. The 36% increase in recall value shows that the model maintains a strong precision-recall balance. This proves our method to be reliable without overfitting or excessive false positives. These results suggest that quantum kernels can capture complex data correlations that classical kernels may overlook.

5.5 Conclusion and Future Work

Our findings demonstrate that effective preprocessing combined with appropriate quantum feature map design enables accurate quantum-based anomaly detection, even in complex physiological datasets. These results highlight the potential of quantum machine learning for stress-related applications in older adults using wearable sensors, underscoring its promise for future smart healthcare systems. However, quantum kernel computation remains resource-intensive on current NISQ hardware, making real-time classification challenging due to limited qubit coherence and execution constraints. Addressing these limitations will require continued research into computational optimizations such as circuit parallelization [151, 152], hybrid and ensemble methods [153], and emerging architectures such as quantum

neural networks [154]. Further study on noise behavior and mitigation techniques is also essential, given the susceptibility of NISQ devices to hardware-induced errors.

This chapter also shows that emotional states can be inferred solely from physiological signals, eliminating the need for intrusive facial recognition technologies. While classical machine learning models provide solid performance for estimating emotion intensities, our hybrid quantum SVM consistently achieved higher accuracy and robustness across all emotion classes, even with relatively small datasets. These findings demonstrate the value of integrating quantum-enhanced kernels with wearable sensor data to build privacy-conscious emotion recognition systems. The implications are especially relevant for individuals with Alzheimer's Disease and Related Dementias (ADRD), veterans with Post-Traumatic Stress Disorder (PTSD), and others who face cognitive or emotional challenges. The proposed methodology supports passive and unobtrusive monitoring in clinical, residential, and assisted living environments. Future work should investigate advanced QML architectures and explore real-time deployment opportunities through optimized edge-quantum workflows.

Chapter 6

Detection of Physiological Data Tampering Attacks with Quantum Machine Learning

This chapter examines the use of quantum-hybrid machine learning models to detect tampering or anomalous alterations in physiological data streams. Data integrity is a critical concern in health-monitoring systems, particularly when measurements inform medical decisions or support older adults who rely on remote monitoring technologies. The chapter demonstrates how quantum fidelity kernels and quantum-enhanced feature embeddings can amplify subtle inconsistencies that arise when physiological signals are manipulated or otherwise deviate from expected patterns. These quantum-based representations provide a powerful mechanism for distinguishing genuine data from corrupted or adversarial input.

Beyond presenting the anomaly detection framework, the chapter discusses the broader implications for cyber-physical health systems, including wearable monitoring platforms used in aging populations. Because older adults may depend on continuous physiological sensing for health assessments, ensuring the trustworthiness and security of their data is essential. By evaluating quantum kernel–based One-Class SVMs

and comparing them with classical baselines, this chapter highlights the potential of quantum machine learning to enhance data reliability and improve safety in real-world health applications.

6.1 Recent Works and Motivation

In contemporary healthcare, wearable and implanted devices are becoming widespread for monitoring physiological parameters. These devices frequently connect to cloud-based platforms, enabling faster decision-making. However, the increased reliance on networked systems has prompted significant concerns regarding data integrity, particularly in the context of cyber risks such as tampering attempts. Data tampering, in which an adversary manipulates physiological data, can lead to inaccurate diagnoses and potentially hazardous medical decisions. If these systems get compromised, it could lead to disastrous incidents involving health and safety [155]. Unattended devices are vulnerable to physical harm by hackers, leading to data alteration and invasions. In most cases, hackers can modify software, install malware, swap out the device for a malicious clone, control sensors, change hardware, and more [156]. A compromised sensor could transmit erroneous information to the gateway node, resulting in inaccurate diagnoses and treatments that could endanger the patient [157]. Machine learning has long been used to solve such tampering detection problems [158, 159, 160]. However, with the increasing complexities of datasets and ML models, classical algorithms are starting to fall short. Again, the adversary or hackers benefit from easily available resources to further complicate the detection process [161, 162].

Now with emerging technologies like Quantum Machine Learning (QML), the problems can be re-evaluated from different perspectives. It uses quantum computing concepts to improve existing machine learning approaches and can detect subtle patterns in tampered data that may bypass traditional methods [60, 61]. Due to their enhanced capacity for pattern recognition, quantum algorithms like quantum support

vector machines have the potential to greatly improve the detection of anomalies in a dataset. Cultice *et al.* [129] showed the advantages of QML in anomaly detection in cyber-physical systems. However, the performance of QML in physiological data tampering detection is still an under-explored domain with several open research questions.

6.2 Contributions

To address these questions, this chapter explores the application of hybrid QML in detecting tampering attacks on physiological data. We also compare the effectiveness of QML-based approaches with classical machine learning techniques. Our contributions are as follows:

- We experimented with three types of physiological sensor datasets with one, two, and three class labels respectively.
- We explored the performance of a threat model that performs two types of white box attacks on physiological sensor data.
- We utilize a quantum hybrid framework for the detection model that uses kernel-based SVM with a quantum kernel circuit.
- We report a detection accuracy of around 75% to 95% for data poisoning attacks for our hybrid QML framework. We also show the comparison of performance in detection between classical and quantum machine learning models.

Based on the contemporary literature, this is the first work in this domain that utilizes quantum machine learning to detect physiological data tampering. We have mentioned our motivation and contributions in Section 6.1. The rest of this chapter has the following structure: Section 6.3 discusses briefly the threat model and the attack. Section 6.4 describes the framework of our hybrid QML detection model and its training. Section 6.5 contains information about our experimental result.

Section 6.6 wraps up the report by suggesting potential future directions for this research.

6.3 Threat Model

A threat model defines the risk conditions and their associated danger distributions [163]. Here, we define risk conditions based on the likelihood and severity of adversarial tampering in cloud-connected physiological monitoring systems. The primary risk conditions include unauthorized data access, model poisoning via malicious training data, and hardware-level vulnerabilities. Their danger distributions describe how different adversarial attacks impact system integrity. For example, minor statistical deviations in physiological data can go undetected and cause misclassification without triggering system-wide alerts. Our threat model includes two forms of white box attacks: *Data Poisoning* and *Adversarial Perturbation*. According to the model, the adversary acquires unauthorized access to the data stream while it is being transferred to the cloud. This allows them to initiate Adversarial Perturbation and execute a Label Flipping attack. Thus, the model trained on adversarial examples will degrade in performance. We will now briefly describe the attacks.

6.3.1 Data Poisoning

Data poisoning attacks occur when an adversary intentionally manipulates the training data of a machine learning model, deteriorating its performance. Label flipping, a typical method of data poisoning, involves an attacker changing the labels of specific training samples while leaving their features unaltered. In targeted poisoning, the attacker attempts to trick the model into creating specific mistakes, usually against particular instances or classes.

6.3.2 Adversarial Perturbation

An adversarial perturbation attack takes advantage of the vulnerability of machine learning models by introducing small, carefully constructed perturbations to input data, causing the model to make inaccurate predictions [164]. The attacker aims to generate an adversarial example $x = x + \delta$, where δ is a small perturbation added to the original input x, such that: $f_\theta(x') \neq f_\theta$.

6.4 Proposed Tampering Detection Method

Our tampering detection approach uses Quantum Hybrid One-Class SVM to address data tampering as an anomaly detection problem. The detection model proceeds in three steps. In the classical domain, it first performs both preprocessing and feature reduction on the data sent from the threat model. Next, a quantum machine calculates quantum fidelity from the kernel. Lastly, the classical domain is used to optimize the One-Class SVM. Figure 6.1 illustrates our proposed method's overall workflow. This method's SVM-based design utilizes quantum fidelity to map the feature space and identify similarities in data. We will now describe these steps in detail.

6.4.1 Dataset Preprocessing and Feature Extraction

We selected three types of datasets having one, two, and three labels for our experiment. Table 6.1 shows a summary of their descriptions. RESTINGECG [165] has one class that represents the ECG signal stream in resting conditions. They recorded continuous EEG signals across 72 channels with electrodes. All channels were amplified in 24-bit DC mode, which was initially sampled at 2,048 Hz (400 Hz bandwidth) and subsequently downsampled online to 256 Hz. EEG signals were captured using a common mode sense (CMS) electrode. Half-cell potentials at the electrode/gel/skin interface were kept within 40 mV, following established standards.

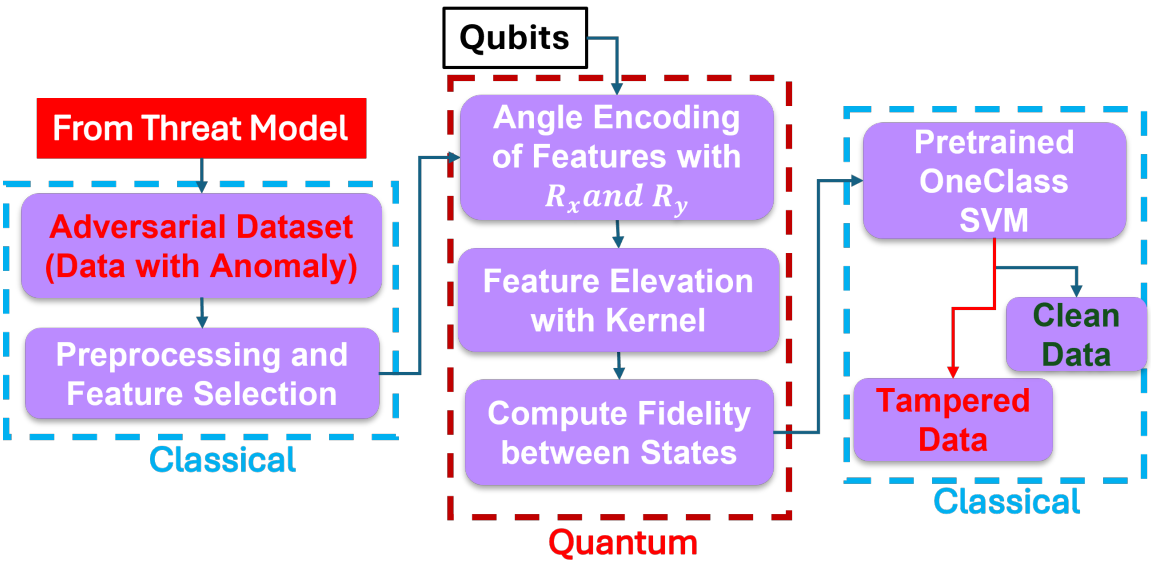


Figure 6.1: Workflow of Proposed Hybrid QML model for Tampering Detection

Table 6.1: Description of Dataset

Dataset	Signals	Class	Participants Age	Extracted Features
RESTINGECG [165]	ECG	One	22 (18-26)	12
EPHNOGRAM [136, 135]	ECG, PCG	Two	24 (23-29)	24
Stress [137, 44, 132]	EDA, BVP, ST, IBI	Three	40 (60-80)	60

The EPHNOGRAM [136, 135] project has two class labels, namely *normal* and *anomaly*. It aimed to develop low-cost, low-power devices capable of simultaneously recording electrocardiogram (ECG) and phonocardiogram (PCG) data and additional channels to capture environmental audio noise. The current dataset, recorded using version 2.1 of the hardware, was collected from 24 healthy persons.

The Stress dataset [44, 132] has three labels: *not-stressed, low-stress* and *high-stress*. Here, 40 older adults aged 60 to 80 were chosen as the target demographic. To provide an unbiased dataset, each subject completed a thorough screening process before participating in the experimental methodology to rule out any pre-existing medical issues. Physiological data were gathered using the Empatica E4 wristband, which includes sensors for skin temperature (ST), photoplethysmography (PPG), and electrodermal activity (EDA) while they participate in TSST (Trier Social Stress Test). Their cortisol samples were collected from their saliva from a time interval to be used as a labeling scheme.

12 statistical features are retrieved from each time domain signal in the EPHNOGRAM and RESTINGECG datasets, making it a total of 24 features for EPHNOGRAM and 12 for RESTINGECG. From the Stress dataset, 60 features were extracted.

6.4.2 Sampling and Feature Reduction

Due to the limitation of resources in terms of memory and runtime, we downsampled the dataset with a moving average window of 60 samples. It helps reduce the noise and also captures the time series patterns. The average samples are calculated using $avg_t = \frac{1}{w} \times \sum_{i=t-w+1}^{t} x_i$. Here, w is the window length and x_i is the sample at time i. Next, the categorical samples are substituted with numerical histogram values that are better suited to the ML algorithms. The dataset is then standardized with Z-score normalization by $x_{std} = \frac{x_o - \mu}{\sigma}$ where, x_0 is the original sample, μ is the sample mean and σ is the standard deviation. This moves the mean of the distribution

to zero and the standard deviation to one. Furthermore, standardization tends to accelerate convergence for optimization techniques such as gradient descent. Again, the computational complexity and runtime depend mainly on the number of qubits. With each number of additional qubits, the complexity increases quadratically. For our experiment, we decided to use 6 qubits that can hold 12 features. Thus, we used PCA (Principal Component Analysis) [166] as a feature reduction technique.

6.4.3 Encoding and Feature Map

To map features into qubits for building kernel circuits, we employed two rotation gates, R_x and R_y. Given two classical features x_1 and x_2, they can be encoded into the quantum state of a qubit using these rotation gates. The R_x gate applies a rotation about the x-axis of the Bloch sphere, mathematically defined as $R_x(\theta) = e^{-i\theta X/2}$, where X is the Pauli-X matrix and θ is the angle of rotation, which can be set as proportional to x_1. Similarly, the R_y gate rotates the qubit around the y-axis and is defined as $R_y(\phi) = e^{-i\phi Y/2}$, with Y being the Pauli-Y matrix and ϕ proportional to x_2. By applying $R_x(x_1)$ followed by $R_y(x_2)$ to each qubit, the two features x_1 and x_2 are mapped to the qubit's quantum state, creating a parameterized quantum state that can be used in quantum kernels. This process encodes classical information into quantum circuits, allowing for quantum-enhanced classical operations. Figure 6.2 shows the quantum circuit that maps the feature into qubits and creates the quantum kernel. The encoded data goes through a series of CNOT (Controlled-NOT) gates. These gates create entanglement between pairs of qubits, which allows the quantum circuit to capture correlations between the data features. Entanglement is a key resource in quantum computation, as it enables the circuit to represent more complex patterns and relationships in the data. After the entangling operations, the circuit applies inverse operations to effectively complete the quantum feature mapping. Following this, the qubits are measured on a computational basis, yielding classical

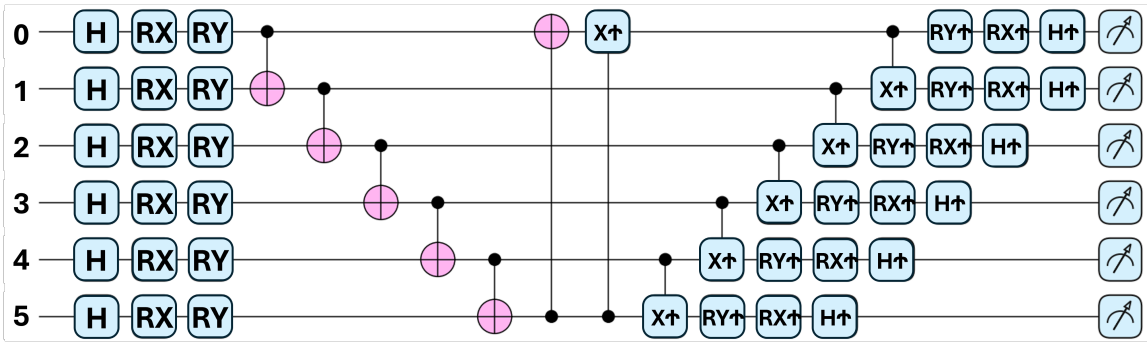
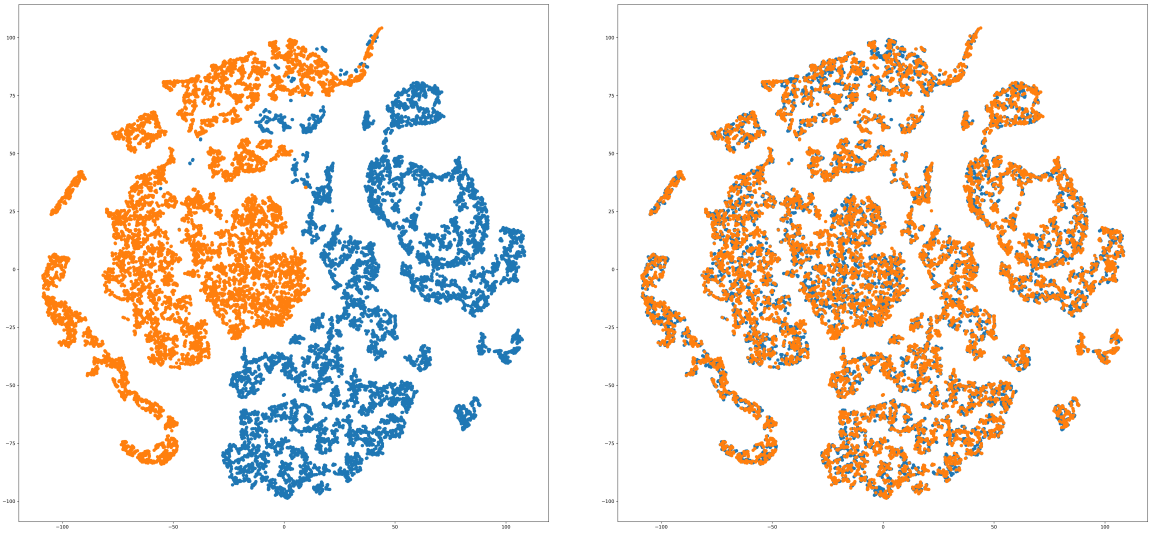


Figure 6.2: Kernel Computation Circuit. $(0-5)$ are the qubits, H is the Hadamard gate, R_x and R_y are the Rotation gates, $\oplus$ denotes the C-NOT gates and † represents the Reverse operations



(a) Clean Dataset before the Attack

(b) Adversarial Dataset after the Attack

Figure 6.3: 2D Projection of EPHNOGRAM Dataset Before and After Data Poisoning Attack. (*blue* and *orange* represent two different classes

data as the output. This output is used to calculate the similarity between two data points in the quantum feature space, which forms the quantum kernel.

6.4.4 Adversarial Attack with Threat Model

We implemented white box attacks on all three datasets and monitored the change in performance for a base model. We see a significant drop in performance in the base model in terms of accuracy, as seen in Table 6.2. The metrics are calculated from the SVM dot product kernel averaging 5-fold of the validation set. As a visual example, Figure 6.3 shows the 2D projection of the EPHNOGRAM dataset before and after the label-flipping attack. The blue and orange markers represent two distinct classes fairly separable in Figure 6.3a but after the attack, they are no longer distinguishable, as seen in Figure 6.3b.

6.4.5 Tampering Detection with One-Class SVM

One-Class SVM works by finding the optimal hyperplane that best separates the samples from potential outliers. Using a kernel can elevate the data to a higher-dimensional feature space, making it easier to draw decision boundaries. The purpose of the training phase is to find the optimal hyperplane with the biggest margin that divides the data points of different classes. In kernel-based SVMs, the optimization problem is addressed under the feature space generated by the kernel function of choice. Additionally, the data points closest to the decision boundary's hyperplane are known as support vectors, which SVM learns during training and are used to define the decision boundary. The driving function for a One-Class SVM can be represented in Eq. (6.1).

$$f(x) = sign\left(\sum_{i=1}^{N}\sum_{j=1}^{N}\alpha_{ij} \times k(x_i, x_j) + b - \rho\right) \tag{6.1}$$

Table 6.2: Average Performance of Dot Product Kernel SVM for 5-fold Data Before and After Adversarial Attack

Dataset	**Class**	**Accuracy(%) (Clean Data)**	**Attack Type**	**Accuracy(% ↓) (Adv Data)**
RESTING ECG	One	85.009	Label Flipping	63.205(↓21.804)
			Adv Perturbation	83.745(↓1.264)
EPHNOGRAM	Two	99.975	Label Flipping	49.556(↓50.419)
			Targeted Poisoning	30.525(↓69.45)
			Adv Perturbation	62.333(↓37.642)
Stress	Three	80.752	Label Flipping	43.878(↓36.874)
			Targeted Poisoning	47.788(↓32.964)
			Adv Perturbation	75.764(↓4.987)

Table 6.3: Detection Accuracy for Test Set of Proposed Framework after 5-fold Cross validation

Method	**Dataset**	**Class**	**Features**	**Attack Type**	**Detection (%)**
Classical	RESTING ECG	One	12	Label Flipping	67.67
				Adv Perturbation	60.52
	EPHNOGRAM	Two	12	Label Flipping	75.01
				Targeted Poisoning	50.36
				Adv Perturbation	68.99
	Stress	Three	12	Label Flipping	80.75
				Targeted Poisoning	65.23
				Adv Perturbation	58.95
Quantum	RESTING ECG	One	12	Label Flipping	75.03
				Adv Perturbation	55.16
	EPHNOGRAM	Two	12	Label Flipping	75.23
				Targeted Poisoning	66.66
				Adv Perturbation	60.09
	Stress	Three	12	Label Flipping	95.11
				Targeted Poisoning	49.96
				Adv Perturbation	45.99

Here, $\{x_i\}_{i=1}^{N}$ is training data from one class, $k(x_i, x_j)$ is the dot product kernel. α_{ij} and b are the constants used as a multiplier and bias term and ρ is the threshold. The optimization is done to maximize the margin between the decision boundary represented as Eq. (6.2) for a given constraint mentioned in Eq. (6.3).

$$\underset{\alpha,\rho}{maximize} = \frac{1}{2} \times ||\omega||^2 - \rho + \frac{1}{\nu \times N} \sum_{i=1}^{N} \zeta_i \tag{6.2}$$

$$when, \omega^T \times \phi(x_i) + b \geq \rho - \zeta_i \tag{6.3}$$

Here, ω is the vector normal to the decision hyperplane, $\phi(x_i)$ is the basis function, ζ is the relaxation value, and ν controls the trade-off between maximizing the margin and controlling the fraction of outliers. This optimization problem is solved iteratively with gradient descent or sequential minimal optimization. In our work, we used the *scikit-learn* module's *svm* class to implement this.

6.5 Result Analysis and Discussion

After the attack, we evaluated the performance of our hybrid model to detect the tampered data points with 5-fold cross-validation and reported the average of the folds. For datasets with a single label (RESTINGECG), the detection accuracy measures the ratio of altered labels or anomalies and the total number of samples. On the other hand, for the datasets with binary (EPHNOGRAM) and multiclass (Stress) labels, the accuracy was calculated class-wise and averaged on five folds. Table 6.3 shows the performance of detection after each type of attack for the three datasets. It also shows the metrics in comparison to the classical version of the framework.

From the table, it is seen that both classical and quantum SVM performed better against label-flipping attacks. In terms of the datasets, the hybrid model performed about 12% better for RESTINGECG single-label data against label-flipping. For EPHNOGRAM with two labels, detection of label-flipping attacks is similar for

both, but the hybrid model performed better against targeted poisoning attacks. Finally, in the Stress dataset with three labels, the proposed hybrid model has 15% more accuracy against label-flipping attacks. Their performance on Adversarial Perturbation is not satisfactory at all. FGSM-based adversarial perturbations are inherently harder to detect, as the subtle perturbation also keeps the statistical properties intact. Thus, both classical and hybrid models perform poorly against them.

6.6 Conclusion and Future Work

In this work, we explored the performance of hybrid QML for tampering detection. Our experiment on 3 types of datasets makes our findings more viable for real-world deployment. Across all datasets, quantum models consistently achieved higher detection rates for label-flipping attacks compared to classical models, with a maximum of 15% increased accuracy. This suggests that quantum algorithms are better suited for identifying data poisoning attacks. While both models struggle with adversarial perturbation attacks, quantum models demonstrate a slight improvement over classical models in terms of detection accuracy. This suggests that quantum algorithms may offer some advantages in handling more sophisticated attacks. This highlights the need for further research to improve model resilience against these types of attacks in the context of physiological data security. One of the research directions would be to explore more robust machine learning models like Quantum Neural Networks (QNN). Again, different kinds of kernel circuits may provide insight into better kernel and fidelity calculation, resulting in better detection accuracy.

Chapter 7

Scalability Analysis of Quantum Models for Stress, Emotion and Data Tampering Detection

This chapter investigates the scalability of quantum machine learning models within the proposed digital biomarker framework for stress, emotion and data tampering detection. Earlier results demonstrate the feasibility and potential advantages of hybrid quantum–classical approaches. The behavior of these models under increasing qubit counts and higher-dimensional encodings remains open for discussion. A systematic evaluation of scalability is therefore necessary to determine how performance changes as model capacity grows. The primary objective of this study is to analyze how expanding the number of qubits affects performance metrics. The observation will be across multiple tasks, including stress detection, emotion recognition, and tampering detection. The qubit count is pushed to a maximum of 24 qubits. Experiments are conducted by a varying qubit counts with controlled circuit structures. Performance trends are evaluated under similar training protocols and hyperparameters to ensure that observed outcomes are unbiased and comparable.

7.1 Computational Infrastructure and Simulation Constraints

All scalability experiments were conducted using classical simulation of quantum circuits, which introduces substantial computational overhead as qubit count increases. Since statevector simulation scales exponentially with the number of qubits, both memory requirements and runtime grow rapidly when evaluating higher-dimensional feature encodings and deeper circuits. Specifically, statevector simulation requires storing 2^n complex amplitudes for an n-qubit system. This eventually results in exponential growth in memory usage and computational cost may also limit the feasible range of qubit counts even on modern HPC systems.

To enable systematic scaling experiments, all models were executed on High Performance Computing (HPC) clusters. The implementation utilized NVIDIA CUDA-Q for both CPU and GPU-accelerated quantum circuit simulation. Despite GPU acceleration, higher-qubit configurations required continuous execution for multiple days due to the exponential growth in Hilbert space and repeated training runs needed for stability analysis. The runtime allocation of 7 days corresponds to the maximum permitted time by the HPC scheduler that ensured completion of long-running simulations without early termination.

A central component of this investigation is the analysis of angle encoding strategies, including one-feature-per-qubit, two-features-per-qubit, and a combination of both. The one-feature-per-qubit approach scales input dimensionality linearly with the number of qubits. It allows higher-resolution representations at the cost of increased computational complexity. The two-features-per-qubit configuration improves information density within a fixed Hilbert space and reduces qubit requirements.

Table 7.1: HPC Resources and Simulation Details for Quantum Model Scalability Experiments

Experiment	**Qubits**	**CPU Cores allocated**	**Features per Qubit**	**Memory allocated**	**Runtime allocated**
Stress Detection	24	128	1	128 GB	7 days
	20	128	2	128 GB	7 days
Emotion Recognition	24	128	1	128 GB	7 days
	20	128	2	256 GB	7 days
Tampering Detection	12	128	1	350 GB	7 days

Table 7.2: Performance Comparison of Stress Detection for High Qubit Configurations

Method	**Features**	**Value**	**Accuracy**	**Precision**	**Recall**	**F1**
Quantum SVM	24 (24-Qubits) 1:1	Max	100%	100%	100%	1.00
		Avg	75%	65%	88%	0.75
	40 (20-Qubits) 2:1	Max	93%	92%	93%	0.93
		Avg	91%	91%	93%	0.91
Classical SVM	24	Max	96%	99%	93%	0.96
		Avg	89%	94%	78%	0.85
	40	Max	96%	99%	93%	0.96
		Avg	89%	94%	78%	0.85

7.2 Scaling Stress Detection with QML for High Qubit Count

To evaluate the scalability of Quantum Machine Learning (QML) in physiological monitoring, we compared the performance of a Quantum Support Vector Machine (QSVM) against a classical SVM across varying qubit counts and feature-to-qubit mapping strategies. The experiments uses 1500 samples with 1200 training and 300 testing for 5-fold cross-validation. The result is summarized in Table 7.2. Here, "1:1" denotes one feature per qubit, while "2:1" represents two features encoded per qubit. This table demonstrate that increasing the Hilbert space dimensionality through higher qubit counts can improve classification performance under certain configurations. The 24-qubit configuration achieved a peak performance of 100% across all metrics in certain folds. But this configuration also showed higher variance compared to the 20-qubit (40 feature) setup. This suggests a potential overfitting and reduced stability at higher Hilbert space dimensionality.

7.2.1 Result Analysis

Several key observations can be drawn from the transition from low-qubit to high-qubit architectures. While low-qubit counts provided competitive results with classical methods, the high-qubit QSVM demonstrates improved capability in handling higher-dimensional feature representations. Increasing the number of qubits enables richer feature embeddings within the Hilbert space, which contributes to enhanced class separability under certain configurations.

The feature-to-qubit mapping strategy plays a critical role in determining model performance and stability. The 2:1 encoding scheme, where two features are mapped to a single qubit, produced more robust and consistent results, achieving an average accuracy of 91%. But the 1:1 mapping showed higher variability despite achieving perfect performance in certain folds. This indicates that dense encoding acts as a form

of implicit regularization. Multiple features can occupy the same quantum space and avoid over-parameterization.

In terms of sensitivity, the QSVM maintained a high recall of 93% on average at higher feature dimensions. This reduction in false negatives is particularly important in stress detection applications, where missed detections can significantly impact the reliability of real-time monitoring systems. However, the observed scaling behavior is not strictly linear and remains highly dependent on encoding strategies. While increasing qubit count can improve peak predictive performance, certain configurations show higher variance. This indicates a trade-off between expressiveness and model stability. These findings further highlight that performance gains at higher qubit counts must be interpreted alongside the associated exponential increase in computational cost.

7.3 Scaling Emotion Recognition with QML for High Qubit Count

In this experiment, we analyzed the impact of qubit scaling on a three-class emotion recognition task (Negative, Neutral, Positive). We compared a dense encoding strategy that encodes 8 features to 4 qubits against 8 features mapped to 8 qubits.

7.3.1 Result Analysis

Table 7.3 shows the transition from 2:1 to a 1:1 feature-to-qubit mapping. The result shows a notable performance improvements for more polarized emotional states. In particular, the 8-qubit configuration achieved near-perfect precision of 0.999 and accuracy of 0.996 for the Positive class. This indicates that the higher-dimensional quantum feature space enables improved separability for strongly expressed emotional patterns. Similarly, recall for the Negative class increased to 0.950. This can reflect enhanced sensitivity to minor physiological variations associated with negative affect.

Table 7.3: Comparison of Emotion Recognition Performance: 4-Qubit vs. 8-Qubit Configurations

Qubits	**Map**	**Negative**				**Neutral**				**Positive**			
		Acc.	**Pre.**	**Rec.**	**F1**	**Acc.**	**Pre.**	**Rec.**	**F1**	**Acc.**	**Pre.**	**Rec.**	**F1**
4 Qubits	2:1	0.815	0.887	0.864	0.872	0.874	0.790	0.864	0.824	0.881	0.896	0.898	0.897
8 Qubits	1:1	0.888	0.890	0.950	0.919	0.801	0.811	0.838	0.825	0.996	0.999	0.990	0.995

Table 7.4: Impact of Qubit Scaling on Data Tampering Detection Accuracy

Features	**Attack Type**	**Map**	**Dataset**	**Detection Acc. (%)**
12	Label Flipping	2:1 (6-Q)	RESTING ECG	75.03
			EPHNOGRAM	75.23
			Stress	95.11
		1:1 (12-Q)	RESTING ECG	85.21
			EPHNOGRAM	55.41
			Stress	94.07

However, these improvements were not uniformly observed across all classes. The Neutral class exhibited a slight decrease in accuracy under the 1:1 mapping. It suggest that intermediate emotional states may occupy a more overlapping or less distinguishable region of the feature space. As a result, increasing Hilbert space dimensionality does not necessarily translate to improved performance for classes with inherently ambiguous boundaries. These findings highlight that the benefits of qubit scaling are class-dependent. While higher qubit counts can significantly enhance discrimination for well-separated classes, they may introduce instability or reduced performance for classes that lie closer to decision boundaries. This further points out the importance of encoding strategies and task-specific characteristics when evaluating scalability in quantum machine learning models.

7.4 Scaling Data Tampering Detection for High Qubit Count

This section evaluates the impact of increasing the qubit-to-feature ratio on the detection of label-flipping attacks across diverse physiological datasets. In contrast to the previous tasks of stress and emotion recognition, tampering detection introduces an adversarial dimension in which the model must identify subtle distribution shifts caused by corrupted labels rather than natural class. This makes the scalability analysis particularly important, as higher-dimensional quantum feature spaces may either enhance or hinder the detection of such attack.

To systematically analyze this effect, we compare two encoding strategies similar to stress and emotion detection. Here for scalability, 12 input features are encoded into 12 qubits. The transition increases the dimensionality of the Hilbert space and allows for a more expressive representation of the input data. However, it also introduces additional complexity that may impact the model's ability to generalize under adversarial conditions. Experiments were conducted across the same physiological

datasets from Chapter 6 to evaluate how different data modalities respond to qubit scaling. The results of this scaling analysis are summarized in Table 7.4.

7.4.1 Result Analysis

The experimental results demonstrate that increasing the qubit-to-feature ratio does not produce uniform improvements across all datasets. Rather it reveals strong domain-dependent behavior. For the RESTING ECG dataset, scaling from a 2:1 mapping to a 1:1 mapping resulted in a notable performance gain of over 10%, achieving a detection accuracy of 85.21%. In contrast, the EPHNOGRAM dataset showed a performance degradation under the 1:1 mapping. The accuracy drops from 75.23% to 55.41%. This indicates that for features present in this dataset, increased qubit dimensionality may introduce unnecessary complexity and reduces the model's ability to capture relevant tampering patterns. In this case, the 2:1 encoding provide a more effective representation.

For the Stress dataset, detection performance remained consistently high across both configurations, with accuracy values around 94%. This stability suggests that the underlying feature representations for stress-related physiological signals are sufficiently separable even at lower qubit counts. At the same time, additional Hilbert space dimensionality does not significantly alter decision boundaries.

These findings indicate that the impact of qubit scaling in tampering detection is highly dependent on the structure and modality of the underlying physiological data. While higher qubit counts can enhance performance for certain signal types, they may also introduce instability or reduced effectiveness for others. This emphasizes the need for domain and data specific encoding strategies when designing scalable quantum machine learning models for adversarial robustness.

7.5 Effect of Scaling For Quantum Machine Learning Framework across Physiological Domains

The experimental evaluations across stress detection, emotion recognition, and data tampering detection collectively demonstrate that scaling quantum feature maps can enhance predictive performance in physiological signal analysis. It also shows that the observed improvements are non-linear and highly task-dependent. For stress detection, increasing qubit counts enabled significant performance gains. With high-qubit configurations QML achieved near-perfect classification in certain cases. This indicates that higher-dimensional quantum representations can effectively capture separable patterns. In emotion recognition, scaling improved discrimination for strongly expressed emotional states. Particularly for the Positive class the improvement was notable. But there was also limited and inconsistent benefits for Neutral emotional state. This highlights the role of class separability in evaluating scalability gains. Finally, the tampering detection results revealed strong domain-specific behavior, where increased qubit dimensionality improved detection accuracy for structured signals such as RESTINGECG. At the same time it showed reduced performance EPHNOGRAM, emphasizing the sensitivity of adversarial detection to encoding strategies and data modality.

These findings indicate that while increasing qubit counts can improve class separability and peak predictive performance, the benefits are strongly influenced by the inherent characteristics of the underlying physiological signals and the chosen encoding scheme. Additionally, these performance gains must be considered alongside the exponential growth in computational cost associated with higher qubit simulations. Effective scalability in quantum machine learning for digital biomarker applications requires careful balancing between model expressiveness, stability, and computational feasibility. These results shows what matters most is the optimized encoding strategies, rather than simply increasing qubit count to achieve robust, generalizable, and efficient quantum machine learning model for physiological domain.

Chapter 8

Conclusion and Future Research Direction

This dissertation presents an integrated computational and practical framework for developing digital biomarkers of stress, emotion, and related psychosocial states in older adults using wearable physiological data, with an emphasis on continuous and non-intrusive monitoring. Across the dissertation, we established methods for detecting and forecasting stress, estimating emotion intensity, and monitoring physiological data integrity using classical machine learning, deep learning, and hybrid quantum machine learning approaches.

The work began with context-aware stress detection that combines multimodal physiological sensing, such as electrodermal activity, blood volume pulse, heart rate variability, acceleration, and skin temperature, with contextual information to support real-time inference and improved robustness. We further demonstrated that sequential deep learning models, including RNN and LSTM architectures, can capture temporal dependencies in physiological responses and enable forecasting of future stress trajectories and support the broader goal of proactive intervention. In parallel, we studied emotion intensity estimation through regression-based modeling to capture

subtle variations that do not rely on intrusive modalities like images or videos during inference.

To address nonlinear and high-dimensional physiological relationships, we incorporated hybrid quantum-classical learning. Using quantum kernels and parameterized quantum circuits as feature maps, the proposed quantum-hybrid SVM framework aims to improve the separability of overlapping physiological patterns and increase robustness to noise, which is particularly relevant for older-adult physiological variability. The dissertation also extended quantum methods to cybersecurity in healthcare by developing quantum-kernel anomaly detection to identify data tampering and adversarial manipulation in physiological streams. Finally, we investigated the scalability of quantum models by examining qubit scaling and encoding strategies, which concludes that performance gains are not strictly linear with increased qubit counts.

8.1 Future Research Direction

Building on these contributions, a few research directions can further extend the clinical impact, robustness, and implications of the dissertation.

8.1.1 Stress Modeling

Our contribution of stress detection and prediction should extend the current binary or 3 class formulation to multiclass-multilabel stress intensity prediction. Thus the varying levels of stress can be represented more realistically in everyday settings. In addition, stress detection and forecasting can be strengthened by incorporating richer contextual information beyond the lab-task structure such as additional indicators of prior state, recovery patterns, or environmental and activity context to reduce false positives when moving from controlled protocols to continuous real-world

monitoring. Finally, while the current time-series forecasting results support short-horizon prediction using RNN and LSTM architectures and cortisol-referenced ground truth, future studies should investigate longer prediction sequence and explicitly characterize how prediction uncertainty grows with record length.

8.1.2 Emotion Intensity Estimation

The next step is to improve generalization across individuals and by expanding dataset diversity and exploring more expressive modeling approaches. This will include deeper architectures or transfer learning methods and keep the privacy-preserving, wearable-only motivation of the framework intact. Because emotion intensity modeling targets subtle variations rather than coarse categories, future research should also focus on improving calibration and robustness under realistic wearable noise and missingness. These types of noise are expected in continuous monitoring scenarios.

8.1.3 Quantum Machine Learning for Healthcare

An important direction is to move beyond idealized simulation assumptions and evaluate the proposed quantum-kernel and hybrid quantum-classical approaches under realistic Noisy Intermediate-Scale Quantum constraints. This includes studying how encoding strategies, circuit structure, and noise levels affect separability, training stability, and generalization in noisy simulation. Since our work focuses on scaling qubits and we have seen that performance gains are more task-dependent than strictly linear with qubit count, the next step should be to implement the framework in real quantum hardware. In parallel, future work should consider device-level feasibility by measuring runtime, resource requirements, and integration overhead in end-to-end monitoring pipelines.

8.1.4 Scalability

Finally, based on our findings about the task-dependent benefits observed across qubit scaling experiments, future research should explore ensemble strategies that combine classical and quantum models to balance performance and computational cost. Stacking or weighted fusion could use classical models to get stable baseline performance and quantum-kernel learners as supplementary experts in cases where quantum feature mappings make class separability or robustness better.

Bibliography

[1] N. W. Judy, "Microblog: Acute Stress vs. Chronic Stress - Nutrition With Judy," 7 2024, accessed: 2025-11-10. [Online]. Available: https://www.nutritionwithjudy.com/microblog-acute-stress-vs-chronic-stress xiv, 2

[2] L. Rapoliene, A. Razbadauskas, and A. Jurgelėnas, "The reduction of distress using therapeutic geothermal water procedures in a randomized controlled clinical trial," *Advances in preventive medicine*, vol. 2015, p. 749417, 04 2015. xiv, 2, 3

[3] J. K. Hietanen, E. Glerean, R. Hari, and L. Nummenmaa, "Bodily maps of emotions across child development," *Developmental Science*, vol. 19, no. 6, p. 1111–1118, Feb. 2016. [Online]. Available: https://doi.org/10.1111/desc.12389 xiv, 5

[4] N. C. Hass, E. J. S. Schneider, and S.-L. Lim, "Emotional expressions of old faces are perceived as more positive and less negative than young faces in young adults," *Frontiers in Psychology*, vol. 6, Aug. 2015. [Online]. Available: http://dx.doi.org/10.3389/fpsyg.2015.01276 xiv, 4, 5

[5] M. Ávila Villanueva, J. Gómez-Ramírez, F. Maestú, C. Venero, J. Ávila, and M. A. Fernández-Blázquez, "The role of chronic stress as a trigger for the alzheimer disease continuum," *Frontiers in Aging Neuroscience*, vol. 12, 2020. xiv, 15, 16

[6] M. Schuld, A. Bocharov, K. M. Svore, and N. Wiebe, "Circuit-centric quantum classifiers," *Phys. Rev. A*, vol. 101, p. 032308, Mar 2020. xv, 35

[7] B. S. McEwen and P. J. Gianaros, "Central role of the brain in stress and adaptation: Links to socioeconomic status, health, and disease," *Annals of the New York Academy of Sciences*, vol. 1186, no. 1, p. 190–222, Feb. 2010. [Online]. Available: http://dx.doi.org/10.1111/j.1749-6632.2009.05331.x 1, 3

[8] S. Cohen, D. A. Tyrrell, and A. P. Smith, "Psychological stress and susceptibility to the common cold," *New England journal of medicine*, vol. 325, no. 9, pp. 606–612, 1991. 1, 60

[9] A. Keller, K. Litzelman, L. E. Wisk, T. Maddox, E. R. Cheng, P. D. Creswell, and W. P. Witt, "Does the perception that stress affects health matter? the association with health and mortality." *Health psychology*, vol. 31, no. 5, p. 677, 2012. 1, 60

[10] H. Cheng and A. Furnham, "Personality, self-esteem, and demographic predictions of happiness and depression," *Personality and individual differences*, vol. 34, no. 6, pp. 921–942, 2003. 1, 60

[11] S. M. Rothman and M. P. Mattson, "Adverse stress, hippocampal networks, and alzheimer's disease," *NeuroMolecular Medicine*, vol. 12, no. 1, p. 56–70, Nov. 2009. 1, 15, 60

[12] J. P. Aquin, R. El-Gabalawy, T. Sala, and J. Sareen, "Anxiety disorders and general medical conditions: Current research and future directions," *Focus*, vol. 15, no. 2, p. 173–181, Apr. 2017. [Online]. Available: http://dx.doi.org/10.1176/appi.focus.20160044 1

[13] S. Cohen, D. Janicki-Deverts, and G. E. Miller, "Psychological stress and disease," *Jama*, vol. 298, no. 14, pp. 1685–1687, 2007. 1, 3, 82

[14] M. Daly and L. Macchia, "Global trends in emotional distress," *Proceedings of the National Academy of Sciences*, vol. 120, no. 14, Mar. 2023. 1

[15] "What's stressing the stressed? main sources of stress among workers," Oct. 2011, accessed: 2025-11-10. [Online]. Available: https://www150.statcan.gc.ca/n1/pub/11-008-x/2011002/article/11562-eng.htm 3

[16] V. Lorant, "Socioeconomic inequalities in depression: A meta-analysis," *American Journal of Epidemiology*, vol. 157, no. 2, p. 98–112, Jan. 2003. [Online]. Available: http://dx.doi.org/10.1093/aje/kwf182 3

[17] "Osh in figures: Stress at work – facts and figures – european sources online," accessed: 2025-11-10. [Online]. Available: https://www.europeansources.info/record/osh-in-figures-stress-at-work-facts-and-figures/ 3

[18] R. T. Iversen, "The mental health of seafarers," *International Maritime Health*, vol. 63, no. 2, pp. 78 – 89, 2012. [Online]. Available: https://journals.viamedica.pl/international_maritime_health/article/view/26143 3

[19] M. Svetina, M. Perkovič, C. Yang, Y. Gu, A. Mindadze, N. Mikeltadze, L. Davitadze, and G. Gabedava, "Factors impacting seafarers' mental health and career intentions," *INQUIRY: The Journal of Health Care Organization, Provision, and Financing*, vol. 61, Jan. 2024. 3

[20] R. W. Levenson, "Basic emotion questions," *Emotion Review*, vol. 3, no. 4, p. 379–386, Sep. 2011. [Online]. Available: http://dx.doi.org/10.1177/1754073911410743 3

[21] S. D. Kreibig, "Autonomic nervous system activity in emotion: A review," *Biological Psychology*, vol. 84, no. 3, p. 394–421, Jul. 2010. [Online]. Available: http://dx.doi.org/10.1016/j.biopsycho.2010.03.010 3, 4

[22] W. H. O. WHO, "Mental health," Oct. 2025, accessed: 2025-11-10. [Online]. Available: https://www.who.int/news-room/fact-sheets/detail/mental-health-strengthening-our-response 4

[23] S. Taylor, N. Jaques, E. Nosakhare, A. Sano, and R. Picard, "Personalized multitask learning for predicting tomorrow's mood, stress, and health," *IEEE Transactions on Affective Computing*, 2017. 4, 6, 42, 65, 66, 83

[24] N. Jaques, S. Taylor, A. Sano, and R. Picard, "Multi-task multi-kernel learning for estimating individual wellbeing," *Neural Information Processing Systems (NeurIPS) Workshop on Multimodal Machine Learning*, 2015. 4, 6, 42

[25] N. Jaques, S. Taylor, A. Azaria, A. Ghandeharioun, A. Sano, and R. Picard, "Predicting students' happiness from physiology, phone, mobility, and behavioral data," in *2015 International Conference on Affective Computing and Intelligent Interaction (ACII)*, 2015, pp. 222–228. 6, 42

[26] T. Zhang, A. El Ali, C. Wang, A. Hanjalic, and P. Cesar, "Cornet: Fine-grained emotion recognition for video watching using wearable physiological sensors," *Sensors*, vol. 21, no. 1, p. 52, 2020. 6, 71, 85

[27] T. K. Hui and R. S. Sherratt, "Coverage of emotion recognition for common wearable biosensors," *Biosensors*, vol. 8, no. 2, p. 30, 2018. 6, 69, 71

[28] M. B. Harms, "Stress and exploitative decision-making," *Journal of Neuroscience*, vol. 37, no. 42, pp. 10 035–10 037, 2017. 14

[29] G. Giannakakis, D. Grigoriadis, K. Giannakaki, O. Simantiraki, A. Roniotis, and M. Tsiknakis, "Review on psychological stress detection using biosignals," *IEEE Transactions on Affective Computing*, vol. 13, no. 1, pp. 440–460, 2022. 14, 15, 41

[30] L. C. Kourtis, O. B. Regele, J. M. Wright, and G. B. Jones, "Digital biomarkers for alzheimer's disease: the mobile/wearable devices opportunity," *npj Digital Medicine*, vol. 2, no. 1, feb 2019. 14

[31] A. N. Edes and D. E. Crews, "Allostatic load and biological anthropology," *American Journal of Physical Anthropology*, vol. 162, no. S63, p. e23146, 2017. 15

[32] K. Opoku Asare, I. Moshe, Y. Terhorst, J. Vega, S. Hosio, H. Baumeister, L. Pulkki-Råback, and D. Ferreira, "Mood ratings and digital biomarkers from smartphone and wearable data differentiates and predicts depression status: A longitudinal data analysis," *Pervasive and Mobile Computing*, vol. 83, p. 101621, 2022. 15, 38

[33] B. Saylam and o. D. incel, "Quantifying digital biomarkers for well-being: Stress, anxiety, positive and negative affect via wearable devices and their time-based predictions," *Sensors*, vol. 23, no. 21, 2023. 15, 38

[34] Y. Jiang, W. Wang, T. Scargill, M. Rothman, J. Dunn, and M. Gorlatova, "Digital biomarkers reflect stress reduction after augmented reality guided meditation: a feasibility study," in *Proceedings of the 2022 Workshop on Emerging Devices for Digital Biomarkers*, ser. DigiBiom '22. New York, NY, USA: Association for Computing Machinery, 2022, p. 29–34. 15

[35] G. Giannakakis, M. Pediaditis, D. Manousos, E. Kazantzaki, F. Chiarugi, P. Simos, K. Marias, and M. Tsiknakis, "Stress and anxiety detection using facial cues from videos," *Biomedical Signal Processing and Control*, vol. 31, pp. 89–101, 2017. 15, 41

[36] M. S. H. Onim, E. Rhodus, and H. Thapliyal, "A review of context-aware machine learning for stress detection," *IEEE Consumer Electronics Magazine*, pp. 1–6, 2023. 17, 21, 41, 69, 84

[37] M. A. Birkett, "The trier social stress test protocol for inducing psychological stress," *Journal of Visualized Experiments*, no. 56, oct 2011. 17

[38] M. S. H. Onim and H. Thapliyal, "Casd-oa: Context-aware stress detection for older adults with machine learning and cortisol biomarker," in *Proceedings of the Great Lakes Symposium on VLSI 2023*, ser. GLSVLSI '23. New York, NY, USA: Association for Computing Machinery, 2023, p. 103–108. 18, 21, 41, 83

[39] E. K. Rhodus, M. S. H. Onim, C. Roberts, S. Kumar, A. M. Burhan, and H. Thapliyal, "Utilization of wearable devices as means for remote digital biometric data collection in a community-based, rural randomized controlled trial among alzheimer's disease dyads," *Alzheimer's & Dementia*, vol. 20, p. e090883, 2024. 21, 69, 70, 83, 84

[40] I. U. Din, A. Almogren, J. J. Rodrigues, and A. Altameem, "Advancing secure and privacy-preserved decision-making in iot-enabled consumer electronics via multimodal data fusion," *IEEE Transactions on Consumer Electronics*, 2024. 21, 69, 70, 83, 84

[41] A. Badawi, S. Elmoghazy, S. Choudhury, S. Elgazzar, K. Elgazzar, and A. Burhan, "A novel multimodal system to predict agitation in people with dementia within clinical settings: A proof of concept," *arXiv preprint arXiv:2411.08882*, 2024. 21, 69, 70, 71, 83, 84, 85

[42] 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. 65 573–65 590, 2020. 21, 39, 40, 69, 84

[43] 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. 21, 39, 40, 41, 69, 84

[44] M. S. H. Onim, H. Thapliyal, and E. K. Rhodus, "Utilizing machine learning for context-aware digital biomarker of stress in older adults," *Information*, vol. 15, no. 5, p. 274, 2024. 21, 69, 83, 84, 105, 106

[45] 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. 70 171–70 181, 2020. 21, 39, 40, 69, 84

[46] 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 Symp. on Computational Intelligence, Cognitive Algorithms, Mind, and Brain (CCMB)*, 2014, pp. 39–48. 21, 40, 69, 84

[47] M. S. H. Onim and H. Thapliyal, "Predicting stress in older adults with rnn and lstm from time series sensor data and cortisol," in *2024 IEEE Computer Society Annual Symposium on VLSI (ISVLSI)*. IEEE, 2024, pp. 300–306. 21

[48] M. Belk, D. Portugal, P. Germanakos, J. Quintas, E. Christodoulou, and G. Samaras, "A computer mouse for stress identification of older adults at work," in *User Modeling, Adaptation, and Personalization*. UMAP (extended proceedings), 2016. 21, 40, 41

[49] iMotions A/S, "Facial expression analysis (fea)," 2024, accessed: 2025-03-22. [Online]. Available: https://imotions.com/blog/publications/ 23

[50] K. A. Tychola, T. Kalampokas, and G. A. Papakostas, "Quantum machine learning—an overview," *Electronics*, vol. 12, no. 11, 2023. [Online]. Available: https://www.mdpi.com/2079-9292/12/11/2379 29

[51] J. Li, Y. Li, J. Song, J. Zhang, and S. Zhang, "Quantum support vector machine for classifying noisy data," *IEEE Transactions on Computers*, vol. 73, no. 9, pp. 2233–2247, 2024. 29

[52] F. V. Massoli, L. Vadicamo, G. Amato, and F. Falchi, "A leap among quantum computing and quantum neural networks: A survey," *ACM Comput. Surv.*, vol. 55, no. 5, Dec. 2022. [Online]. Available: https://doi.org/10.1145/3529756 29

[53] M. Kjaergaard, M. E. Schwartz, J. Braumüller, P. Krantz, J. I.-J. Wang, S. Gustavsson, and W. D. Oliver, "Superconducting qubits: Current state of play," *Annual Review of Condensed Matter Physics*, vol. 11, no. Volume 11, 2020, pp. 369–395, 2020. [Online]. Available: https://www.annualreviews.org/content/journals/10.1146/annurev-conmatphys-031119-050605 29

[54] C. Monroe, "Ionq quantum computers: clear to scale," in *APS March Meeting Abstracts*, vol. 2021, 2021, pp. P10–002. 29, 33

[55] R. LaRose and B. Coyle, "Robust data encodings for quantum classifiers," *Phys. Rev. A*, vol. 102, p. 032420, Sep 2020. 31, 90

[56] S. Resch and U. R. Karpuzcu, "Benchmarking quantum computers and the impact of quantum noise," *ACM Comput. Surv.*, vol. 54, no. 7, Jul. 2021. [Online]. Available: https://doi.org/10.1145/3464420 31

[57] A. Javadi-Abhari, M. Treinish, and K. Krsulich, "Quantum computing with Qiskit," 2024. 33

[58] Z. Cai, R. Babbush, S. C. Benjamin, S. Endo, W. J. Huggins, Y. Li, J. R. McClean, and T. E. O'Brien, "Quantum error mitigation," *Rev. Mod. Phys.*, vol. 95, p. 045005, Dec 2023. [Online]. Available: https://link.aps.org/doi/10.1103/RevModPhys.95.045005 33

[59] J. Clark and H. Thapliyal, "Peephole optimization for quantum approximate synthesis," in *2024 25th International Symposium on Quality Electronic Design (ISQED)*, 2024, pp. 1–8. 33

[60] M. Schuld and N. Killoran, "Quantum machine learning in feature hilbert spaces," *Phys. Rev. Lett.*, vol. 122, p. 040504, Feb 2019. 34, 82, 101

[61] V. Havlíček, A. D. Córcoles, K. Temme, A. W. Harrow, A. Kandala, J. M. Chow, and J. M. Gambetta, "Supervised learning with quantum-enhanced feature spaces," *Nature*, vol. 567, no. 7747, pp. 209–212, Mar 2019. 34, 36, 101

[62] C. Cortes and V. Vapnik, "Support-vector networks," *Machine Learning*, 1995. 34

[63] B. E. Boser, I. M. Guyon, and V. N. Vapnik, "A training algorithm for optimal margin classifiers," in *Proceedings of the fifth annual workshop on Computational learning theory*, 1992, pp. 144–152. 34

[64] A. Jafari, A. Ganesan, C. S. K. Thalisetty, V. Sivasubramanian, T. Oates, and T. Mohsenin, "Sensornet: A scalable and low-power deep convolutional neural network for multimodal data classification," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 66, no. 1, 2019. 39

[65] S. Aristizabal, K. Byun, N. Wood, A. F. Mullan, P. M. Porter, C. Campanella, A. Jamrozik, I. Z. Nenadic, and B. A. Bauer, "The feasibility of wearable and self-report stress detection measures in a semi-controlled lab environment," *IEEE Access*, vol. 9, pp. 102 053–102 068, 2021. 39

[66] M. M. Hassan, M. G. R. Alam, M. Z. Uddin, S. Huda, A. Almogren, and G. Fortino, "Human emotion recognition using deep belief network architecture," *Information Fusion*, vol. 51, pp. 10–18, 2019. 39

[67] Siddharth, T.-P. Jung, and T. J. Sejnowski, "Utilizing deep learning towards multi-modal bio-sensing and vision-based affective computing," *IEEE Transactions on Affective Computing*, vol. 13, no. 1, pp. 96–107, 2022. 39

[68] 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. 39, 40

[69] K. Adeli, V. Higgins, M. Nieuwesteeg, J. E. Raizman, Y. Chen, S. L. Wong, and D. Blais, "Biochemical Marker Reference Values across Pediatric, Adult, and Geriatric Ages: Establishment of Robust Pediatric and Adult Reference Intervals on the Basis of the Canadian Health Measures Survey," *Clinical Chemistry*, vol. 61, no. 8, pp. 1049–1062, 08 2015. 41

[70] L. Deng and D. Yu, "Deep learning: Methods and applications," *Found. Trends Signal Process.*, vol. 7, pp. 197–387, 2014. 41

[71] Y. R. Tabar and U. Halici, "A novel deep learning approach for classification of EEG motor imagery signals," *Journal of Neural Engineering*, vol. 14, no. 1, p. 016003, nov 2016. 41

[72] X. Zhai, B. Jelfs, R. H. M. Chan, and C. Tin, "Self-recalibrating surface EMG pattern recognition for neuroprosthesis control based on convolutional neural network," *Frontiers in Neuroscience*, vol. 11, jul 2017. 41

[73] W. Geng, Y. Du, W. Jin, W. Wei, Y. Hu, and J. Li, "Gesture recognition by instantaneous surface EMG images," *Scientific Reports*, vol. 6, no. 1, nov 2016. 41

[74] J. T. Ruiz, J. D. B. Pérez, and J. R. B. Blázquez, "Arrhythmia detection using convolutional neural models," in *Distributed Computing and Artificial Intelligence, 15th International Conference*. Springer International Publishing, jul 2018, pp. 120–127. 41

[75] Y. Xiang, Z. Lin, and J. Meng, "Automatic QRS complex detection using two-level convolutional neural network," *BioMedical Engineering OnLine*, vol. 17, no. 1, jan 2018. 41

[76] R. D. Labati, E. Muñoz, V. Piuri, R. Sassi, and F. Scotti, "Deep-ECG: Convolutional neural networks for ECG biometric recognition," *Pattern Recognition Letters*, vol. 126, pp. 78–85, sep 2019. 41

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

[78] H. Thapliyal, V. Khalus, and C. Labrado, "Stress detection and management: A survey of wearable smart health devices," *IEEE Cons. Electron Mag.*, vol. 6, no. 4, pp. 64–69, 2017. 41

[79] M. U. Ahmed, "A case-based multi-modal clinical system for stress management," Ph.D. dissertation, Mälardalen University, 2010. 41

[80] M. Singh and A. B. Queyam, "A novel method of stress detection using physiological measurements of automobile drivers," *International Journal of Electronics Engineering*, vol. 5, no. 2, pp. 13–20, 2013. 41

[81] D. MacLean, A. Roseway, and M. Czerwinski, "Moodwings: a wearable biofeedback device for real-time stress intervention," in *Proceedings of the 6th International Conference on PErvasive Technologies Related to Assistive Environments*, ser. PETRA '13. New York, NY, USA: Association for Computing Machinery, 2013. 43

[82] U. S. Reddy, A. V. Thota, and A. Dharun, "Machine learning techniques for stress prediction in working employees," in *2018 IEEE International Conference on Computational Intelligence and Computing Research (ICCIC)*, 2018, pp. 1–4. 42, 43, 65, 66

[83] R. Alharthi, R. Alharthi, B. Guthier, and A. El Saddik, "Casp: context-aware stress prediction system," *Multimedia Tools and Applications*, vol. 78, no. 7, pp. 9011–9031, 2019. 42

[84] M. Parent, A. Tiwari, I. Albuquerque, J.-F. Gagnon, D. Lafond, S. Tremblay, and T. H. Falk, "A multimodal approach to improve the robustness of physiological stress prediction during physical activity," in *2019 IEEE International Conference on Systems, Man and Cybernetics (SMC)*, 2019, pp. 4131–4136. 42, 65, 66, 83

[85] H. Yu, E. B. Klerman, R. W. Picard, and A. Sano, "Personalized wellbeing prediction using behavioral, physiological and weather data," in *2019 IEEE EMBS International Conference on Biomedical & Health Informatics (BHI)*, 2019, pp. 1–4. 42, 65, 66

[86] F. D. Martino and F. Delmastro, "High-resolution physiological stress prediction models based on ensemble learning and recurrent neural networks," in *2020 IEEE Symposium on Computers and Communications (ISCC)*, 2020, pp. 1–6. 42, 65, 66

[87] T. Umematsu, A. Sano, S. Taylor, M. Tsujikawa, and R. W. Picard, "Forecasting stress, mood, and health from daytime physiology in office workers and students," in *2020 42nd annual international conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*. IEEE, 2020, pp. 5953–5957. 42, 66

[88] R. Dai, C. Lu, L. Yun, E. Lenze, M. Avidan, and T. Kannampallil, "Comparing stress prediction models using smartwatch physiological signals and participant self-reports," *Computer Methods and Programs in Biomedicine*, vol. 208, p. 106207, 2021. 42, 43, 65, 66

[89] A. Kene and S. Thakare, "Mental stress level prediction and classification based on machine learning," in *2021 Smart Technologies, Communication and Robotics (STCR)*, 2021, pp. 1–7. 42, 44, 65, 66

[90] R. Wang, F. Chen, Z. Chen, T. Li, G. Harari, S. Tignor, X. Zhou, D. Ben-Zeev, and A. T. Campbell, "Studentlife: assessing mental health, academic performance and behavioral trends of college students using smartphones," in *Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, ser. UbiComp '14. New York, NY, USA: Association for Computing Machinery, 2014, p. 3–14. 43

[91] D. Ben-Zeev, E. A. Scherer, R. Wang, H. Xie, and A. T. Campbell, "Next-generation psychiatric assessment: Using smartphone sensors to monitor behavior and mental health." *Psychiatric Rehabilitation Journal*, vol. 38, no. 3, p. 218–226, Sep. 2015. 43

[92] D. Becker, V. Bremer, B. Funk, J. Asselbergs, H. Riper, and J. Ruwaard, "How to predict mood? delving into features of smartphone-based data," *Healthcare Informatics and Information Technology (SIGHealth)*, 2016. 43

[93] A. Pratap, D. C. Atkins, B. N. Renn, M. J. Tanana, S. D. Mooney, J. A. Anguera, and P. A. Areán, "The accuracy of passive phone sensors in predicting daily mood," *Depression and Anxiety*, vol. 36, no. 1, pp. 72–81, 2019. 43

[94] S. Saeb, E. G. Lattie, S. M. Schueller, K. P. Kording, and D. C. Mohr, "The relationship between mobile phone location sensor data and depressive symptom severity," *PeerJ*, vol. 4, p. e2537, Sep. 2016. 43

[95] M. Thelwall, "Tensistrength: Stress and relaxation magnitude detection for social media texts," *Information Processing & Management*, vol. 53, no. 1, pp. 106–121, 2017. 43

[96] H. Lin, J. Jia, J. Qiu, Y. Zhang, G. Shen, L. Xie, J. Tang, L. Feng, and T.-S. Chua, "Detecting stress based on social interactions in social networks," *IEEE Transactions on Knowledge and Data Engineering*, vol. 29, no. 9, pp. 1820–1833, 2017. 43

[97] P. B. Pankajavalli, G. S. Karthick, and R. Sakthivel, "An efficient machine learning framework for stress prediction via sensor integrated keyboard data," *IEEE Access*, vol. 9, pp. 95 023–95 035, 2021. 43

[98] T. Stütz, T. Kowar, M. Kager, M. Tiefengrabner, M. Stuppner, J. Blechert, F. H. Wilhelm, and S. Ginzinger, "Smartphone based stress prediction," in *User Modeling, Adaptation and Personalization: 23rd International Conference, UMAP 2015, Dublin, Ireland, June 29–July 3, 2015. Proceedings 23.* Springer, 2015, pp. 240–251. 43

[99] D. Sharma, N. Kapoor, and S. S. Kang, "Stress prediction of students using machine learning," *International Journal of Mechanical and Production Engineering Research and Development*, vol. 10, no. 3, 2020. 44

[100] C. Ding, Y. Zhang, and T. Ding, "A systematic hybrid machine learning approach for stress prediction," *PeerJ Computer Science*, vol. 9, p. e1154, Feb. 2023. 44

[101] Y. Haque, R. S. Zawad, C. S. A. Rony, H. Al Banna, T. Ghosh, M. S. Kaiser, and M. Mahmud, "State-of-the-art of stress prediction from heart rate variability using artificial intelligence," *Cognit. Comput.*, vol. 16, no. 2, pp. 455–481, Mar. 2024. 44

[102] M. Peruzzini, F. Grandi, and M. Pellicciari, "How to analyse the workers' experience in integrated product-process design," *Journal of Industrial Information Integration*, vol. 12, pp. 31–46, 2018. 44

[103] G. Blandino, "How to measure stress in smart and intelligent manufacturing systems: A systematic review," *Systems*, vol. 11, no. 4, p. 167, 2023. 44

[104] 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 J. Biomed. Health Inform.*, vol. 14, no. 2, pp. 410–417, 2009. 47

[105] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg *et al.*, "Scikit-learn: Machine learning in python," *The Journal of machine Learning research*, vol. 12, pp. 2825–2830, 2011. 47

[106] 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, p. 34, 1998. 47

[107] M. G. Kendall, "A new measure of rank correlation," *Biometrika*, vol. 30, no. 1/2, pp. 81–93, 1938. 51

[108] E. Timotius and G. S. Octavius, "Stress at the workplace and its impacts on productivity: A systematic review from industrial engineering, management, and medical perspective," *Industrial Engineering & Management Systems*, vol. 21, no. 2, pp. 192–205, 2022. 60

[109] Y. Wei, Y. Wu, and J. Tudor, "A real-time wearable emotion detection headband based on eeg measurement," *Sensors and Actuators A: Physical*, vol. 263, pp. 614–621, 2017. 69, 71

[110] M. Ragot, N. Martin, S. Em, N. Pallamin, and J.-M. Diverrez, "Emotion recognition using physiological signals: laboratory vs. wearable sensors," in *Advances in Human Factors in Wearable Technologies and Game Design: Proceedings of the AHFE 2017 International Conference on Advances in Human*

Factors and Wearable Technologies, July 17-21, 2017, The Westin Bonaventure Hotel, Los Angeles, California, USA 8. Springer, 2018, pp. 15–22. 69, 71

[111] H. Kim, J. Ben-Othman, L. Mokdad, and P. Bellavista, "A virtual emotion detection architecture with two-way enabled delay bound toward evolutional emotion-based iot services," *IEEE Transactions on Mobile Computing*, vol. 21, no. 4, pp. 1172–1181, 2020. 69, 71

[112] A. Costa, J. A. Rincon, C. Carrascosa, V. Julian, and P. Novais, "Emotions detection on an ambient intelligent system using wearable devices," *Future Generation Computer Systems*, vol. 92, pp. 479–489, 2019. 69, 71

[113] X. Jiang, X. Zhang, and D. Wu, "Active learning for black-box adversarial attacks in eeg-based brain-computer interfaces," in *2019 IEEE Symposium Series on Computational Intelligence (SSCI)*. IEEE, 2019, pp. 361–368. 69, 71

[114] S. Saganowski, J. Komoszyńska, M. Behnke, B. Perz, D. Kunc, B. Klich, Ł. D. Kaczmarek, and P. Kazienko, "Emognition dataset: emotion recognition with self-reports, facial expressions, and physiology using wearables," *Scientific data*, vol. 9, no. 1, p. 158, 2022. 70, 84

[115] P. Schmidt, A. Reiss, R. Duerichen, C. Marberger, and K. Van Laerhoven, "Introducing wesad, a multimodal dataset for wearable stress and affect detection," in *Proceedings of the 20th ACM International Conference on Multimodal Interaction*, ser. ICMI '18. New York, NY, USA: Association for Computing Machinery, 2018, p. 400–408. 70, 84

[116] H. Z. Wijasena, R. Ferdiana, and S. Wibirama, "A survey of emotion recognition using physiological signal in wearable devices," in *International Conference on Artificial Intelligence and Mechatronics Systems*. IEEE, 2021, pp. 1–6. 70, 84

[117] S. Ba and X. Hu, "Measuring emotions in education using wearable devices: A systematic review," *Computers & Education*, vol. 200, p. 104797, 2023. 70, 84

[118] J. C. Quiroz, E. Geangu, and M. H. Yong, "Emotion recognition using smart watch sensor data: Mixed-design study," *JMIR mental health*, vol. 5, no. 3, p. e10153, 2018. 71

[119] J. S. Heinisch, I. Hübener, and K. David, "The impact of physical activities on the physiological response to emotions," in *2018 IEEE International Conference on Pervasive Computing and Communications Workshops (PerCom Workshops)*. IEEE, 2018, pp. 824–829. 71

[120] C. Xiefeng, Y. Wang, S. Dai, P. Zhao, and Q. Liu, "Heart sound signals can be used for emotion recognition," *Scientific reports*, vol. 9, no. 1, p. 6486, 2019. 71

[121] L. Shu, Y. Yu, W. Chen, H. Hua, Q. Li, J. Jin, and X. Xu, "Wearable emotion recognition using heart rate data from a smart bracelet," *Sensors*, vol. 20, no. 3, p. 718, 2020. 71, 85

[122] J. A. Miranda, M. F. Canabal, L. Gutierrez-Martin, J. M. Lanza-Gutierrez, M. Portela-Garcia, and C. Lopez-Ongil, "Fear recognition for women using a reduced set of physiological signals," *Sensors*, vol. 21, no. 5, p. 1587, 2021. 71, 85

[123] S. Chen, K. Jiang, H. Hu, H. Kuang, J. Yang, J. Luo, X. Chen, and Y. Li, "Emotion recognition based on skin potential signals with a portable wireless device," *Sensors*, vol. 21, no. 3, p. 1018, 2021. 71, 85

[124] B. S. McEwen, "Central effects of stress hormones in health and disease: Understanding the protective and damaging effects of stress and stress mediators," *European journal of pharmacology*, vol. 583, no. 2-3, pp. 174–185, 2008. 82

[125] C. Ciliberto, M. Herbster, A. D. Ialongo, M. Pontil, A. Rocchetto, S. Severini, and L. Wossnig, "Quantum machine learning: a classical perspective," *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, vol. 474, no. 2209, p. 20170551, 2018. 82

[126] N. Moll, P. Barkoutsos, L. S. Bishop, J. M. Chow, A. Cross, D. J. Egger, S. Filipp, A. Fuhrer, J. M. Gambetta, M. Ganzhorn *et al.*, "Quantum optimization using variational algorithms on near-term quantum devices," *Quantum Science and Technology*, vol. 3, no. 3, p. 030503, 2018. 82

[127] P. Rebentrost, M. Mohseni, and S. Lloyd, "Quantum support vector machine for big data classification," *Physical review letters*, vol. 113, no. 13, p. 130503, 2014. 82

[128] S. Y.-C. Chen, C.-H. H. Yang, J. Qi, P.-Y. Chen, X. Ma, and H.-S. Goan, "Variational quantum circuits for deep reinforcement learning," *IEEE Access*, vol. 8, pp. 141 007–141 024, 2020. 82

[129] T. Cultice, M. S. H. Onim, A. Giani, and H. Thapliyal, "Anomaly detection for real-world cyber-physical security using quantum hybrid support vector machines," in *2024 IEEE Computer Society Annual Symposium on VLSI (ISVLSI)*. IEEE, 2024, pp. 619–624. 82, 102

[130] A. Padha and A. Sahoo, "Maqml: a meta-approach to quantum machine learning with accentuated sample variations for unobtrusive mental health monitoring," *Quantum Machine Intelligence*, vol. 5, no. 1, p. 17, 2023. 83

[131] U. Ullah and B. Garcia-Zapirain, "Quantum machine learning revolution in healthcare: a systematic review of emerging perspectives and applications," *IEEE Access*, 2024. 83

[132] M. S. H. Onim and H. Thapliyal, "Predicting stress in older adults with rnn and lstm from time series sensor data and cortisol," in *2024 IEEE Computer Society Annual Symposium on VLSI (ISVLSI)*, 2024, pp. 300–306. 83, 105, 106

[133] H.-K. Shin, W. Lee, J.-H. Yun, and H. Kim, *HAI 1.0: HIL-Based Augmented ICS Security Dataset*. USA: USENIX Association, 2020. 83

[134] H.-K. Shin, W. Lee, J.-H. Yun, and B.-G. Min, "Two ics security datasets and anomaly detection contest on the hil-based augmented ics testbed," in *Cyber Security Experimentation and Test Workshop*, ser. CSET '21. New York, NY, USA: Association for Computing Machinery, 2021, p. 36–40. 83

[135] A. L. Goldberger, L. A. N. Amaral, L. Glass, J. M. Hausdorff, P. C. Ivanov, R. G. Mark, J. E. Mietus, G. B. Moody, C.-K. Peng, and H. E. Stanley, "Physiobank, physiotoolkit, and physionet: Components of a new research resource for complex physiologic signals," *Circulation*, vol. 101, no. 23, Jun. 2000. 83, 105, 106

[136] A. Kazemnejad, P. Gordany, and R. Sameni, "Ephnogram: A simultaneous electrocardiogram and phonocardiogram database," 2021. 83, 105, 106

[137] 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, 2022. 83, 105

[138] N.-D. Mai, H. Barki, and W.-Y. Chung, "A wearable bte-eeg embedded device for emotion monitoring with quantum machine learning," in *2024 Tenth International Conference on Communications and Electronics (ICCE)*, 2024, pp. 562–566. 85, 86

[139] J. Singh, F. Ali, B. Shah, K. S. Bhangu, and D. Kwak, "Emotion quantification using variational quantum state fidelity estimation," *IEEE Access*, vol. 10, pp. 115 108–115 119, 2022. 85, 86

[140] A. A. Chandanwala, S. Bhowmik, P. Chaudhury, and S. C. Pravin, "Hybrid quantum deep learning model for emotion detection using raw eeg signal analysis," 2024. 86

[141] R. Golchha, M. Sahu, and V. Bhateja, "Quantum-based deep learning method for recognition of facial expressions," *Neural Computing and Applications*, mar 2025. 85, 86

[142] J. Singh, K. S. Bhangu, A. Alkhanifer, A. A. AlZubi, and F. Ali, "Quantum neural networks for multimodal sentiment, emotion, and sarcasm analysis," *Alexandria Engineering Journal*, vol. 124, pp. 170–187, 2025. 85, 86

[143] S. Suraj Jai Krishna, M. Anish, A. M. Posonia, J. Albert Mayan, and P. Asha, "Gesture and emotion detection using quantum computing for enhanced recognition and analysis," in *2024 International Conference on Expert Clouds and Applications (ICOECA)*, 2024, pp. 530–535. 86

[144] S. Alsubai, A. Alqahtani, A. Alanazi, M. Sha, and A. Gumaei, "Facial emotion recognition using deep quantum and advanced transfer learning mechanism," *Frontiers in Computational Neuroscience*, vol. Volume 18 - 2024, 2024. 86

[145] Q. Li, D. Gkoumas, A. Sordoni, J.-Y. Nie, and M. Melucci, "Quantum-inspired neural network for conversational emotion recognition," *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 35, no. 15, pp. 13 270–13 278, May 2021. 85

[146] Z. Li, Y. Zhou, Y. Liu, F. Zhu, C. Yang, and S. Hu, "QAP: A quantum-inspired adaptive-priority-learning model for multimodal emotion recognition," in *Findings of the Association for Computational Linguistics: ACL 2023*, A. Rogers, J. Boyd-Graber, and N. Okazaki, Eds. Toronto, Canada: Association for Computational Linguistics, Jul. 2023, pp. 12 191–12 204. 85

[147] Y. Liu, Y. Zhang, and D. Song, "A quantum probability driven framework for joint multi-modal sarcasm, sentiment and emotion analysis," *IEEE Transactions on Affective Computing*, vol. 15, no. 1, pp. 326–341, 2024. 85

[148] K. A. Woźniak, V. Belis, E. Puljak, P. Barkoutsos, G. Dissertori, M. Grossi, M. Pierini, F. Reiter, I. Tavernelli, and S. Vallecorsa, "Quantum anomaly detection in the latent space of proton collision events at the lhc," 2023. 90

[149] V. Bergholm, J. Izaac, M. Schuld, C. Gogolin, S. Ahmed, V. Ajith, M. S. Alam, G. Alonso-Linaje, B. AkashNarayanan, A. Asadi *et al.*, "Pennylane: Automatic differentiation of hybrid quantum-classical computations," *arXiv preprint arXiv:1811.04968*, 2018. 91

[150] Qiskit contributors, "Qiskit: An open-source framework for quantum computing," 2023. 91

[151] Y. Ohkura, T. Satoh, and R. Van Meter, "Simultaneous execution of quantum circuits on current and near-future nisq systems," *IEEE Transactions on Quantum Engineering*, vol. 3, pp. 1–10, 2022. 98

[152] C. Domeniconi and D. Gunopulos, "Incremental support vector machine construction," in *Proceedings 2001 IEEE International Conference on Data Mining*, 2001, pp. 589–592. 98

[153] D. B. Araya, K. Grolinger, H. F. ElYamany, M. A. Capretz, and G. Bitsuamlak, "An ensemble learning framework for anomaly detection in building energy consumption," *Energy and Buildings*, vol. 144, pp. 191–206, 2017. 98

[154] A. Kukliansky, M. Orescanin, C. Bollmann, and T. Huffmire, "Network anomaly detection using quantum neural networks on noisy quantum computers," *IEEE Transactions on Quantum Engineering*, vol. 5, pp. 1–11, 2024. 99

[155] J. Malik, R. Muthalagu, and P. M. Pawar, "A systematic review of adversarial machine learning attacks, defensive controls, and technologies," *IEEE Access*, vol. 12, pp. 99 382–99 421, 2024. 101

[156] M. Li, W. Lou, and K. Ren, "Data security and privacy in wireless body area networks," *IEEE Wireless Communications*, vol. 17, no. 1, pp. 51–58, 2010. 101

[157] T. Dimitriou and K. Ioannis, "Security issues in biomedical wireless sensor networks," in *2008 First International Symposium on Applied Sciences on Biomedical and Communication Technologies*, 2008, pp. 1–5. 101

[158] A. K. Pathak, S. Saguna, K. Mitra, and C. Åhlund, "Anomaly detection using machine learning to discover sensor tampering in iot systems," in *ICC 2021 - IEEE International Conference on Communications*, 2021, pp. 1–6. 101

[159] J. Pacheco and S. Hariri, "Anomaly behavior analysis for iot sensors," *Transactions on Emerging Telecommunications Technologies*, vol. 29, no. 4, p. e3188, 2018, e3188 ETT-16-0338.R1. 101

[160] P. D. Rosero-Montalvo, Z. István, P. Tözün, and W. Hernandez, "Hybrid anomaly detection model on trusted iot devices," *IEEE Internet of Things Journal*, vol. 10, no. 12, pp. 10 959–10 969, 2023. 101

[161] G. Sun, Y. Cong, J. Dong, Q. Wang, L. Lyu, and J. Liu, "Data poisoning attacks on federated machine learning," *IEEE Internet of Things Journal*, vol. 9, no. 13, pp. 11 365–11 375, 2022. 101

[162] F. A. Yerlikaya and Şerif Bahtiyar, "Data poisoning attacks against machine learning algorithms," *Expert Systems with Applications*, vol. 208, p. 118101, 2022. 101

[163] D. Xu, M. Tu, M. Sanford, L. Thomas, D. Woodraska, and W. Xu, "Automated security test generation with formal threat models," *IEEE Transactions on Dependable and Secure Computing*, vol. 9, no. 4, pp. 526–540, 2012. 103

[164] S.-M. Moosavi-Dezfooli, A. Fawzi, O. Fawzi, and P. Frossard, "Universal adversarial perturbations," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 1765–1773. 104

[165] L. T. Trujillo, C. T. Stanfield, and R. D. Vela, "The effect of electroencephalogram (eeg) reference choice on information-theoretic measures of the complexity and integration of eeg signals," *Frontiers in Neuroscience*, vol. 11, 2017. 104, 105

[166] F. L. Gewers, G. R. Ferreira, H. F. D. Arruda, F. N. Silva, C. H. Comin, D. R. Amancio, and L. D. F. Costa, "Principal component analysis: A natural approach to data exploration," *ACM Comput. Surv.*, vol. 54, no. 4, May 2021. 107

Vita

Md. Saif Hassan Onim was born in December 1997 in Mymensingh, Bangladesh. He earned his Bachelor of Science in Electrical, Electronic and Communication Engineering (EECE) from the Military Institute of Science and Technology (MIST), Dhaka, Bangladesh. After graduation, he worked as an Adjunct Lecturer in EECE, MIST for 2 and a half years.

In Fall 2022 he joined as a Ph.D. student in Computer Engineering under the supervision of Professor Himanshu Thapliyal at the University of Tennessee, Knoxville. His research focuses on machine learning and quantum machine learning for digital health, with emphasis on wearable physiological sensing, time-series modeling, and hybrid quantum–classical methods for stress, emotion, and Alzheimer's-related biomarker detection. He has experience building end-to-end research pipelines, including data acquisition and processing, reproducible HPC workflows, and deployment-oriented implementations on embedded and FPGA platforms.

He has published and presented in IEEE, ACM, and biomedical computing venues and has contributed professional service through IEEE leadership roles. He was the recipient of the prestigious UT-ORII GATE Fellowship, and he was also awarded the Gonzalez Family Award for Outstanding Graduate Teaching Assistant in Computer Engineering at UTK. He is currently a member of the IEEE and ACM. Upon graduation, Saif intends to work in academia to further contribute to the teaching and research community.