الجمهورية الجزائرية الديمقراطية الشعبية

Ministry of Higher Education and Scientific Research

Yahia Fares University of Medea

Faculty of Sciences

Department of Mathematics and Computer Science



Graduation Project Report

Submitted for Master of Computer Science Degree

Speciality: Engineering and technologies of web systems

Development of a Real-Time Stress Detection System

Proposed and directed by: Dr. Ahmed KHELDOUN

Prepared and presented by: Mr.Abdelkader HASNAOUI

Defended on June, 2025, In front of the jury composed of:

Mr. Mohamed Yassine HAYI – Yahia Fares University – *President*

Ms. Nesrine BOUDEKHNANI – Yahia Fares University – *Examiner*

2024-2025

Acknowledgements

First and foremost,
I am deeply thankful and grateful to ***ALLAH*** *for providing me with the patience and strength to persevere throughout this work.*
I like to express my profound thanks to my supervisor,
Dr.Ahmed KHELDOUN*, for his essential patience, support, and direction during the project.*
Furthermore, I would want to express my gratitude to all of the instructors from whom I have gained knowledge during the last five years.
In revising several sections of this thesis,
QuillBot was helpful,
and I am grateful for that.

Dedication

I dedicate this work to ***ALLAH*** *that gave me the power and the strength to keep going until this day .*
To my first supporters in life my parents
who were always here for me and by my side I love you
mom *and* ***dad*** *.*
To my dear ***brother*** *and* ***sister****,*
whose love, support, and encouragement
have been a constant source of strength and inspiration
throughout my academic journey.
To my friends.
to everyone believed in me
and to me. . .

Abstract

This research presents the development of a real-time stress detection system based on physiological signals, with a particular focus on Photoplethysmography (PPG) signals derived from the WESAD dataset. The goal is to accurately classify stress levels using a combination of traditional machine learning models (SVM, Random Forest, XGBoost) and deep learning approaches (CNN, LSTM, and hybrid architectures). The system involves multiple stages, including signal preprocessing, statistical feature extraction, data augmentation, and model optimization.

Experimental results show that the LSTM model achieved the best performance, reaching a classification accuracy of over 93.23% and a macro F1-score of 0.92, while maintaining a low prediction latency of approximately 0.06 seconds per sample. These results confirm the suitability of the proposed system for real-time applications. Overall, this method provides a foundation for non-invasive, continuous stress monitoring in real-world settings.

Keywords: Stress Detection, Real-Time Monitoring, Photoplethysmography PPG(BVP), Deep Learning, Machine Learning, WESAD Dataset.

Ce projet présente l'élaboration d'un système de détection du stress en temps réel, reposant sur l'analyse de signaux physiologiques, en particulier la photopléthysmographie (PPG). L'objectif est de catégoriser avec précision les niveaux de stress en recourant à un ensemble de modèles d'apprentissage automatique classiques (SVM, Forêt Aléatoire, XGBoost) ainsi qu'à des approches d'apprentissage profond (CNN, LSTM et modèles hybrides). Le système comprend plusieurs étapes, telles que le prétraitement des signaux, l'extraction de caractéristiques statistiques, l'augmentation des données et l'optimisation des hyperparamètres.

Les résultats expérimentaux montrent que lRésumée modèle LSTM a surpassé les autres configurations, atteignant une précision de 93,23% et un score F1 macro de 0,92, avec un temps de prédiction moyen d'environ 0,06 seconde par échantillon, ce qui rend le système adapté à une utilisation en temps réel. Ce dispositif constitue ainsi une base prometteuse pour la surveillance continue et non intrusive du stress dans des contextes réels.

Mots-clés : Détection de stress, surveillance en temps réel, photopléthysmographie PPG (BVP), apprentissage profond, apprentissage automatique, ensemble de données WESAD.

ملخص

يعرض هذا المشروع تطوير نظام لاكتشاف التوتر (القلق) في الوقت الحقيقي، يعتمد على تحليل الإشارات الفسيولوجية، وخصوصًا إشارة التخطيط الضوئي للدم. يهدف هذا النظام إلى تصنيف مستويات التوتر بدقة من خلال استخدام مجموعة من نماذج التعلم الآلي التقليدية (مثل الغابة العشوائية...)، إلى جانب أساليب التعلم العميق (مثل النماذج الهجينة).

يتكوّن النظام من عدة مراحل تشمل: المعالجة المسبقة للإشارات، استخراج الخصائص الإحصائية، زيادة البيانات، وتحسين إعدادات النماذج (الضبط الفائق للمعلمات).

تُظهر النتائج التجريبية أن نموذج الشّبكة العصبونيّة ذات الذّاكرة الطّويلة قصيرة المدى قد تفوّق على باقي التكوينات، حيث حقق دقة قدرها ٢٣.٩٣ %، ودرجة بلغت ٩٢.٠، مع متوسط زمن توقّع يقارب ٠.٦٠ ثانية لكل عينة، مما يجعل النظام مناسبًا للاستخدام في الوقت الحقيقي. يُعتبر هذا النظام بذلك قاعدة واعدة للمراقبة المستمرة وغير التداخلية لمستويات التوتر في البيئات الحقيقية.

الكلمات المفتاحية: اكتشاف التوتر، المراقبة في الوقت الحقيقي، التخطيط الضوئي للدم ، التعلم العميق، التعلم الآلي، مجموعة بيانات .

Contents

list of tables **8**

list of figures **10**

List of Abbreviations **11**

1 Introduction **12**
1.1 Problematic . . . 13
1.2 Motivation and Objectives . . . 14
1.3 Thesis Structure . . . 14

2 Machine Learning Overview **15**
2.1 Introduction . . . 16
2.2 Artificial Intelligence . . . 16
2.3 Machine Learning . . . 17
2.3.1 Supervised Learning . . . 17
2.3.2 Unsupervidsed . . . 21
2.3.3 Semi-supervised . . . 21
2.3.4 Reinforcement . . . 21
2.4 Deep Learning . . . 22
2.4.1 Convolutional neural network . . . 22
2.4.2 Long Short Term Memory . . . 26
2.5 Ensemble learning . . . 27
2.5.1 Types of Ensemble Models . . . 27
2.5.2 Ensemble learning techniques . . . 27
2.6 Conclusion . . . 29

3 Stress Overview **30**
3.1 Introduction . . . 31
3.2 Signs and Symptoms of Stress Overload . . . 32
3.3 Biosignals related to stress . . . 33
3.3.1 Heart Rate (HR) . . . 34
3.3.2 Electroencephalogram (EEG) . . . 36
3.3.3 Electrodermal Activity (EDA) . . . 38
3.3.4 Respiration and Breath Rate (BR) . . . 40
3.3.5 Photoplethysmography (PPG) . . . 42
3.4 Related Work . . . 44

3.5 Conclusion 45

4 Conception **46**
4.1 Indroduction 47
4.2 Suggested Approach 47
4.3 Phases 48
4.3.1 Preprocessing 48
4.3.2 Data Augmentation 51
4.3.3 LSTM Model for Stress Classification 52
4.4 Evaluation metrics 54
4.5 Conclusion 55

5 Results And Discussion **56**
5.1 Indroduction 57
5.2 Development environment 57
5.2.1 Hardware specifications 57
5.2.2 Environment and Development Tools : 57
5.3 Dataset for stress detection using PPG signals 60
5.3.1 Results : 62
5.4 Comparison of Proposed Model with Related Works 68
5.5 Conclusion 69

General Conclusion **70**

List of Tables

3.1 Signs and Symptoms of Stress Overload 32
3.2 Main Features Related to stress detection for ECG signal 36

4.1 LSTM Model Architecture 53
4.2 Summary of the proposed LSTM-based attention model architecture. . 53

5.1 Comprehensive Comparison of Machine Learning and Deep Learning Classifiers 66
5.2 Comparison of Proposed Model with Related Works in rPPG-based Stress Detection 68

List of Figures

2.1 Artificial Intelligence Components 16
2.2 Machine learning techniques 17
2.3 Supervised Learning 18
2.4 DecisionTree 19
2.5 Random forest 20
2.6 Unsupervised Learning 21
2.7 CNN architecture. 23
2.8 Convolution Layer Without Padding 24
2.9 convolution operation 24
2.10 Illustration of Pooling Techniques 25
2.11 LSTM basic unit 26

3.1 Schematic schematic illustrating typical locations of wearable sensors on the human body 34
3.2 Typical ECG signal with characteristics peaks P,Q,R,S,T 35
3.3 Typical raw EEG acquired with biosignalplux 38
3.4 Typical components of EDA signal. 39
3.5 Typical raw EDA acquired with biosignalplux 40
3.6 Raw Respiration signal acquired with biosignalplux 41
3.7 Raw PPG signal and its corresponding ECG. 42

4.1 Schematic Representation Development of a Real-Time Stress Detection System. 48
4.2 Original PPG (BVP)Signal (Before Windowing) 49
4.3 PPG(BVP) Signal with Sliding Windows (After Windowing) 50

5.1 Label Distribution in WESAD Data 61
5.2 Label Distribution in WESAD Data After Augmentation 61
5.3 Example PPG (BVP) Signals from WESAD Subjects 62
5.4 LSTM Model Training and Validation Performance Over Epochs 62
5.5 Confusion Matrix of LSTM Predictions for Baseline, Stress, and Amusement Classes 63
5.6 LSTM with attention Model Training and Validation Performance Over Epochs 64
5.7 Confusion Matrix - LSTM with Attention 65
5.8 Confusion Matrix of CNN Predictions for Baseline, Stress, and Amusement Classes 76

5.9 CNN Model Training and Validation Performance Over Epochs 76
5.10 Confusion Matrix - Ensemble (CNN + LSTM) 77
5.11 Training Validation Metrics per Epoch 77

List of Abbreviations

ACC	Accelerometer
AI	Artificial Intelligence
BVP	Blood Volume Pulse
CNN	Convolutional Neural Network
DL	Deep Learning
DT	Decision Tree
ECG	Electrocardiogram
EDA	Electrodermal Activity
EEG	Electroencephalogram
EMG	Electromyography
HRV	Heart Rate Variability
KNN	K-Nearest Neighbors
LSTM	Long Short-Term Memory
ML	Machine Learning
PPG	Photoplethysmography
RF	Random Forest
rPPG	Remote Photoplethysmography
SVM	Support Vector Machine
WESAD	Wearable Stress and Affect Detection Dataset
XGBoost	Extreme Gradient Boosting

Chapter 1

Introduction

Systems for stress detection and management have garnered considerable attention in recent years. Their roots are rooted in clinical psychology and medical diagnostics, originally formulated to evaluate mental health risks in controlled settings. In the last ten years, progress in sensor technology and machine learning has broadened their applicability outside conventional clinical environments. These systems now cater to several sectors, including workplace wellness, healthcare, education, and personal fitness, allowing real-time monitoring of physiological biomarkers (e.g., heart rate variability, cortisol levels) and behavioral indicators (e.g., speech patterns, facial expressions).

Their rapid advancement has transformed the approach of people and companies towards mental well-being. With the ongoing advancement of wearable devices and AI-driven analytics, stress detection systems are set to become intricately woven into everyday life. This increase presents issues, especially with data privacy, algorithmic precision, and the ethical ramifications of ongoing biometric monitoring. Maintaining the security of sensitive health data is crucial, since breaches may reveal highly personal information. The increasing need for non-invasive, real-time stress monitoring requires strong frameworks to align innovation with ethical limitations.

Subtle or intense stress reactions complicate these issues. Classifying stress in real-time, particularly in dynamic or loud contexts, is essential for facilitating appropriate treatments and reducing long-term health concerns. To overcome these challenges and facilitate the secure, fair incorporation of stress detection technologies into society, the capacity to consistently recognize and understand stress signals, even in their nascent phases, is essential.

1.1 Problematic

Contemporary methods for stress identification and treatment in healthcare and business environments are often reactive and disjointed. The dependence on subjective self-reporting instruments, postponed professional evaluations, or irregular monitoring presents considerable obstacles for people and organizations aiming to proactively manage stress. Technical constraints, like the absence of real-time physiological data integration, together with ethical issues around privacy in continuous biometric monitoring, further hinder adoption. Uncontrolled stress presents significant dangers, including prolonged mental health decline, diminished productivity, and physical health issues such as cardiovascular strain. Although current frameworks emphasize practicality and privacy protection, they often fail to reconcile accuracy with prompt response. This creates a significant void: devices capable of real-time stress detection while upholding ethical norms. Regulatory and social obstacles, including data security and reducing invasiveness, reflect the intricacy of incorporating these technology into everyday life. Confronting these tensions—leveraging innovation to enhance well-being while safeguarding privacy and autonomy—continues to be a critical task for both academics and governments.

1.2 Motivation and Objectives

The aim of this project is to provide real-time solutions for stress detection and categorization to reduce dangers linked to uncontrolled stress, including its physiological and psychological effects. These systems provide the continuous monitoring of stress biomarkers and behavioral indicators, therefore enabling healthcare practitioners, companies, and people to proactively tackle stress-related issues, boost mental well-being, and improve quality of life via prompt interventions.

1.3 Thesis Structure

This work will contain four chapters

- **Chapter 1 :** This chapter provides a comprehensive overview of the current status of artificial intelligence (AI) approaches used for stress detection and categorization.
- **Chapter 2 :** This chapter examined methodologies for real-time stress detection, emphasizing physiological and behavioral signs to assess mental well-being. It tackled issues in precisely detecting stress reactions and examined techniques to improve system dependability and performance.
- **Chapter 3 :** This chapter delves into the underlying workings of our system that detects stress. It shows how to build a system in a systematic way, from organizing data to training detection models.
- **Chapter 4 :** This section discusses the technical attributes of our project, including the development environment, employed datasets, and implementation details.

Chapter 2

Machine Learning Overview

2.1 Introduction

New technologies have emerged in recent years with the potential to revolutionize our daily lives, work environments, and interactions with the environment. Among them, three terms stand out: Artificial Intelligence (AI), Machine Learning (ML), and Deep Learning (DL). While they are often used interchangeably, they are not the same thing. We will explore what AI, ML, and DL are and what separates them.

2.2 Artificial Intelligence

Artificial intelligence is a broad field of research. Real-world problems can be solved with the help of this field of study. Artificial intelligence is the process of making a computer, a robot controlled by a computer, or software think intelligently in a way that is comparable to that of intelligent humans. This is achieved through the study of human cognition, learning, decision-making, and problem-solving processes, and then applying the findings to the development of intelligent software and systems. A range of deep learning and machine learning approaches are used to achieve this, each of which is appropriate for a specific set of objectives and tasks.
Figure 2.1 illustrates these components, providing a visualization of the foundational elements of AI.

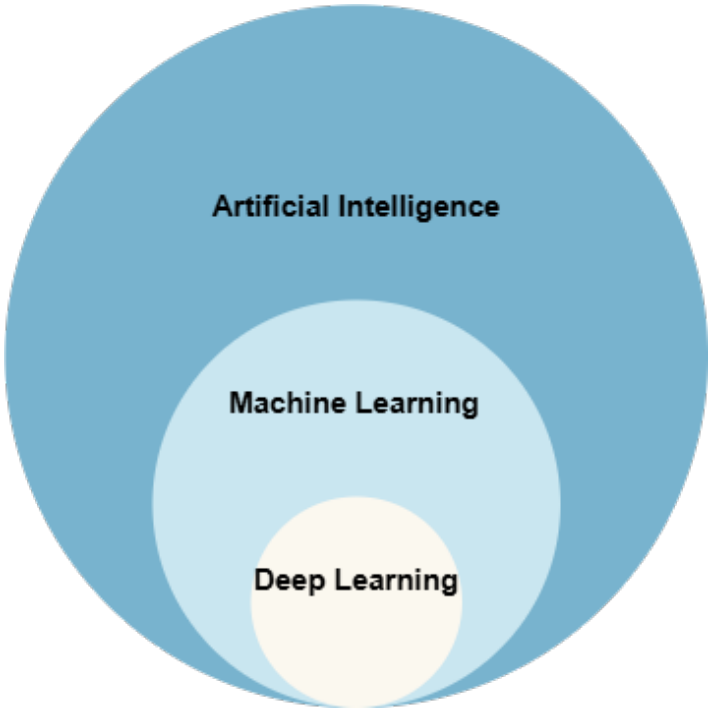


Figure 2.1: Artificial Intelligence Components

2.3 Machine Learning

The ability for computers to learn to recognize patterns in data without human intervention is known as machine learning. Statistical modeling and algorithm development that make predictions or educated judgments based on past data is at the heart of this field. There is a wide variety of approaches within the discipline, all with its own unique purpose and set of difficulties to solve. Now we will examine the basic machine learning methods shown in Figure 2.2, focusing on their distinct characteristics and real-world applications.

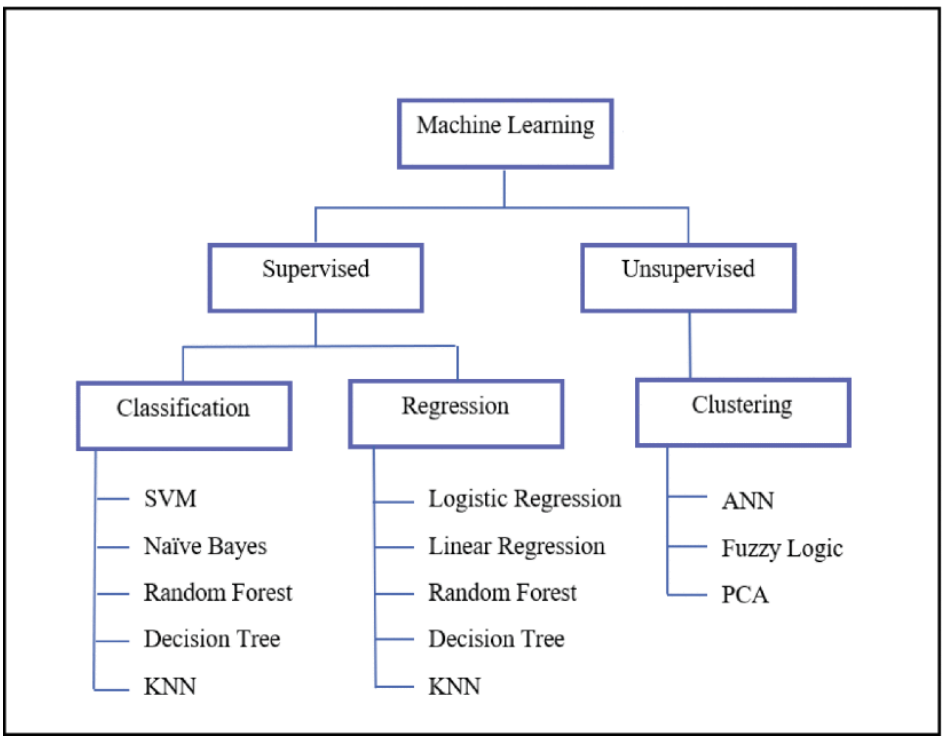


Figure 2.2: Machine learning techniques

2.3.1 Supervised Learning

In the field of machine learning, supervised learning refers to a paradigm in which algorithms are taught to map input data to desired outputs by making use of labeled datasets, which are examples that have been pre-tagged with correct answers or classifications. During the training process, the system utilizes these known input-output pairings to identify patterns, which it then uses construct a predictive model. The algorithm, after it has been trained, is able to generalize this knowledge in order to assess new data that it has not before encountered and to create correct predictions or judgments.
as seen in Figure 2.3, when a decision border splits two sets of data points into two distinct groupings.

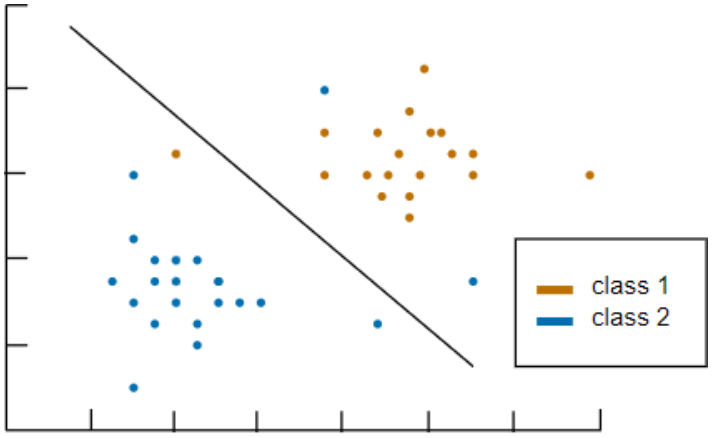


Figure 2.3: Supervised Learning

Support Vector Machines

The goal of the Support Vector Machine (SVM) algorithm is to identify a hyperplane within an N-dimensional space (where N represents the number of features) that effectively distinguishes between data points. Since multiple hyperplanes can separate the two classes, the objective is to select the one that maximizes the margin—the distance between the closest data points of each class. A larger margin enhances the model's ability to classify future data points with greater confidence [15] .

SVM classifiers are divided into linear and nonlinear types depending on how the data is distributed.

- **Linear SVMs:** Applied to datasets that can be separated without transformation. The decision boundary is defined by the equation $wx + b = 0$. Two classification strategies exist:
  - Hard-margin SVMs, which enforce perfect separation without errors.
  - Soft-margin SVMs, which introduce slack variables (ξ) to allow some misclassification. A tuning parameter C controls how strict the boundary is.
- **Nonlinear SVMs:** Designed for data that cannot be linearly separated. They transform data into a higher-dimensional space using kernel functions, such as polynomial, radial basis function (RBF), and sigmoid kernels, enabling more effective classification. The kernel trick helps simplify computation.
  Support Vector Regression (SVR) Support Vector Regression (SVR) adapts SVM principles for regression, focusing on predicting continuous values. Unlike traditional linear regression, SVR does not require predefined relationships between variables, making it well-suited for time series forecasting. This section explores how to build an SVM classifier, compares it with other supervised learning algorithms, and discusses its real-world applications .
  Steps to Build an **SVM Classifier** Data Preparation: The dataset is divided into training and testing sets. Performing exploratory data analysis (EDA) helps identify missing values and outliers. Model Training and Evaluation: Using a machine learning library like scikit-learn, the SVM classifier is trained, and its

accuracy is assessed with F1-score, precision, or recall. Optimizing Performance: Grid search and cross-validation are used to fine-tune hyperparameters like the kernel function, regularization (C), and gamma to enhance accuracy.

Decision Trees

Since a random forest consists of multiple decision trees, it is useful to first understand how a decision tree works. A decision tree begins with a fundamental question, such as "Should I go surfing?" From there, a series of follow-up inquiries like "Is there a long-period swell?" or "Is the wind blowing offshore?".help narrow down the option. These questions represent decision nodes, which divide the data into distinct branches. Each observation follows either a "Yes" or "No" path based on whether it meets the criteria until reaching a final decision, represented by a leaf node. Decision trees try to determine the ideal strategy to split data into meaningful groupings. They are often trained using the Classification and Regression Tree (CART) method, which assesses splits using metrics such as Gini impurity, information gain, or mean squared error (MSE). In this scenario, the decision tree represents a categorization issue, where the alternative outcomes are "surf" or "don't surf."
While decision trees are widely used in supervised learning, they can suffer from issues like bias and overfitting. However, when numerous decision trees are integrated in a random forest model, the overall accuracy increases notably when the individual trees are varied and uncorrelated.

as seen in Figure 2.4 .

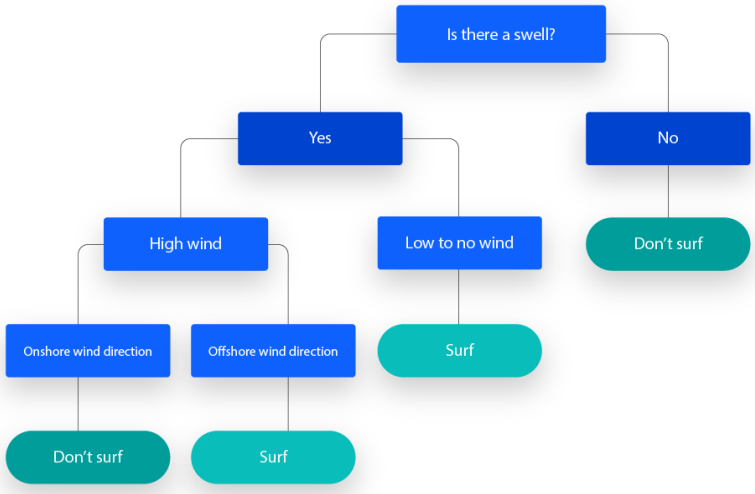


Figure 2.4: DecisionTree

Random forest

The Random Forest algorithm is an ensemble learning method that improves prediction accuracy by combining multiple decision trees. Each tree is trained on a randomly

selected subset of data using bootstrap sampling, and at each split, a random subset of features is considered. This randomness reduces overfitting and enhances model performance. By aggregating the outputs of diverse trees, Random Forest captures different aspects of the data, making it more reliable than individual decision trees. For example, in customer churn prediction, different trees may prioritize factors like payment history or customer service interactions, leading to a more balanced prediction. Similarly, in house price estimation, trees may focus on different property attributes, ensuring a robust and well-rounded outcome.[37]
In the context of the "Should I surf?" scenario, many persons may pose distinct inquiries to reach a choice. The incorporation of variability via feature selection in the random forest model mitigates the hazards of overfitting, bias, and variance, hence enhancing prediction accuracy. Prior to training, three essential hyperparameters must be established: node size, the quantity of trees, and the number of sampled features. Upon configuration, the random forest classifier is capable of managing both classification and regression tasks.
A random forest comprises several decision trees, each trained on a bootstrap sample—a dataset randomly selected from the training set with replacement. Furthermore, one-third of the initial dataset is excluded as the out-of-bag (OOB) sample, subsequently used for validation purposes. Additional randomization is included by feature bagging, augmenting model variety and diminishing correlation among trees. The predictive methodology is contingent upon the nature of the issue. In regression tasks, predictions from several trees are averaged, whereas in classification tasks, the final class is established by majority vote. The out-of-bag sample is used for cross-validation to enhance the model's predictive accuracy.

as seen in Figure 2.5

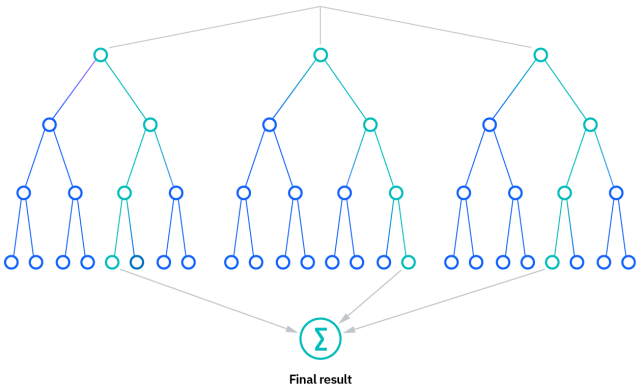


Figure 2.5: Random forest

XGBoost

Extreme Gradient Boosting, often known as XGBoost, is a machine learning toolkit that is open-source and distributed. It makes use of gradient-boosted decision trees, which are a kind of supervised learning technique that use gradient descent. Due to the fact that it is known for its great speed, efficiency, and scalability, it is an excellent choice for dealing with massive datasets. The XGBoost algorithm, which was developed by Tianqi Chen at the University of Washington, is an improved version of the

gradient boosting algorithm that adheres to the same underlying foundation. It does this by integrating weak decision trees into a more strong model via the accumulation of residuals, which in turn has the effect of improving predictive performance.

2.3.2 Unsupervidsed

Unsupervised learning is a method of machine learning that is used to extract insights from data without the use of predetermined labels or categories. It is able to independently discover hidden patterns, relationships, or groups within raw datasets, in contrast to supervised algorithms, which need annotated samples. In this method, the emphasis is placed on discovering internal data patterns via the use of self-organized analysis. This method operates without the need for explicit instructions or predetermined results.
As can be seen in the graphic, data points that have the same color are grouped together together into different clusters.
as seen in Figure 2.6

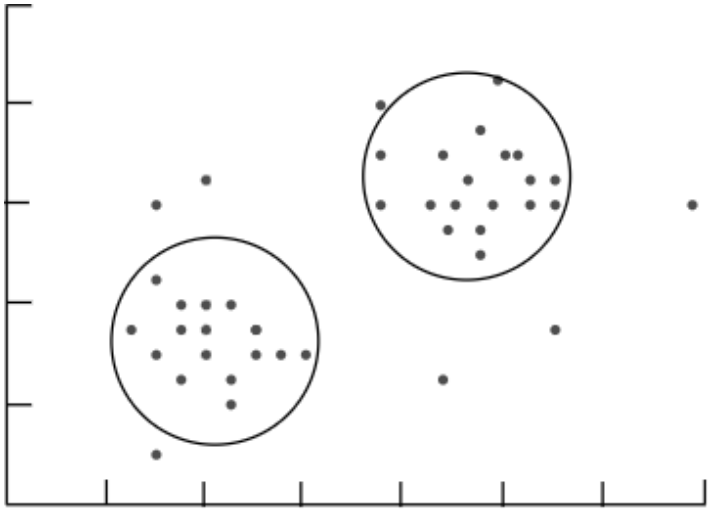


Figure 2.6: Unsupervised Learning

2.3.3 Semi-supervised

Semi-supervised learning bridges the gap between supervised and unsupervised paradigms, combining a limited set of labeled data with abundant unlabeled examples during training. This hybrid approach allows models to infer patterns from unstructured data while refining their accuracy using the smaller annotated dataset. However, unlike supervised learning, the algorithm is trained on a dataset that contains both labeled and unlabeled data. Semi-supervised learning is particularly useful when there is a large amount of unlabeled data available, but it's too expensive or difficult to label all of it. [16]

2.3.4 Reinforcement

Reinforcement learning (RL) is a machine learning technique that instructs software to do certain desired behaviors. Reinforcement learning relies on incentivizing preferred

actions and penalizing undesirable ones.
A reinforcement learning agent, the software entity undergoing training, may sense and understand its surroundings, perform actions, and learn via trial and error.
Reinforcement learning is one of several approaches developers use to train machine learning systems. This approach is important because it empowers an agent to learn to navigate the complexities of the environment for which it was created. For example, an agent can be taught to control a video game, or a robot in an industrial setting can be taught to perform a specific task. Over time, through a feedback system that typically includes rewards and punishments, the agent learns from its environment and optimizes its behavior. [19]

2.4 Deep Learning

Deep learning is a subset of machine learning characterized by its ability to autonomously learn representations from data, such pictures, video, or text, without the need of human domain expertise. The term "deep" in deep learning denotes the several layers of algorithms (or neural networks) that examine data to identify intricate patterns. These architectures are meant to emulate facets of human cognition, acquiring knowledge directly from unrefined, unstructured material without need on manual feature extraction. As the system handles greater information, its hierarchical structure facilitates increasingly precise prediction abilities.

2.4.1 Convolutional neural network

CNN is a deep learning algorithm designed to process image data and perform tasks related to object recognition, such as image classification, detection, and segmentation, utilizing convolutional layers. Each layer functions akin to a magnifying glass. In image processing, convolution refers to a mathematical operation that integrates input data with a filter/kernel to generate a feature map.
Convolutional layers are essential for capturing hierarchical patterns within data. As the network deepens, it learns higher-level features by amalgamating lower-level features identified by the filters in preceding layers. This hierarchical representation enables CNNs to proficiently learn and recognize intricate patterns in visual data, rendering them particularly effective for tasks such as image classification and object detection. Figure 2.7 illustrates the architecture of a typical CNN; let us further explore how these layers collaborate to process and extract features from input data.

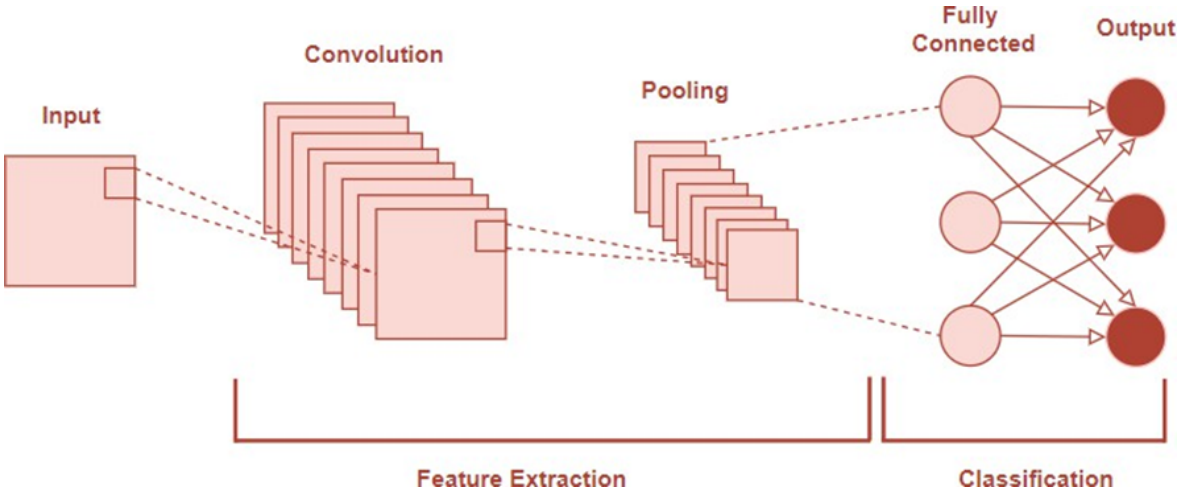


Figure 2.7: CNN architecture.

- **Convolutional layer:** Features are extracted from the input data, which might be photos or any other kind of data, using the convolutional layer. Learnable filters or kernels are part of it. A filter is a tiny matrix that performs a convolution operation on the input data as it slides across it, multiplying each element at each position:

$$x_j = f\left(\sum_{i \in M_j} x_i^{-1} \cdot k_{ij} + b_j\right) \tag{1.1}$$

1. **Padding :** A feature map's output size is affected by three variables during convolution: stride, kernel size, and input feature map size. The output feature map will be smaller than the input if the kernel is applied directly to the input feature map. The supplied feature map's edges may become blank as a consequence. Padding, which is inserting additional pixels around the input feature map's border before to convolution, is used to preserve the boundary information.

   Without Padding The following procedure is used to compute the size of the output matrix without padding:

$$x = n - f + 1 \times n - f + 1 \tag{1.2}$$

   Figure 2.8 depicts a convolutional layer without padding for your visual representation.

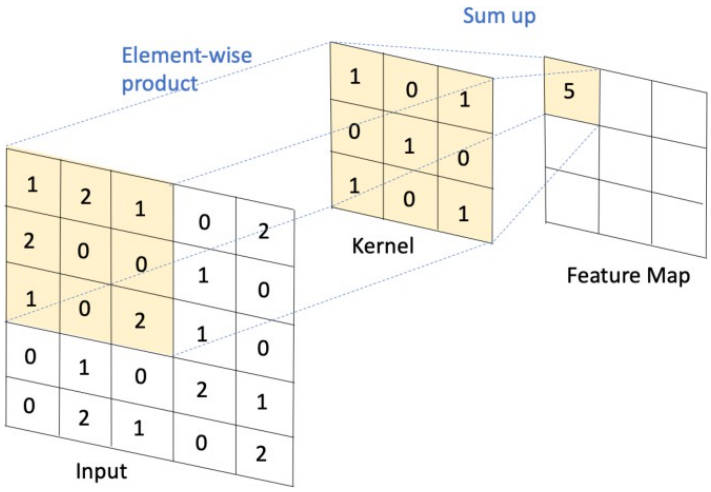


Figure 2.8: Convolution Layer Without Padding

With Padding For the padding scenario, this process determines the size of the output matrix:

$$x = n + 2p - f + 1 \times n + 2p - f + 1 \tag{II.8}$$

as well as the padding value:

$$p = \frac{f-1}{2} \tag{1.3}$$

Check out Figure for a visual representation of a convolutional layer with padding.

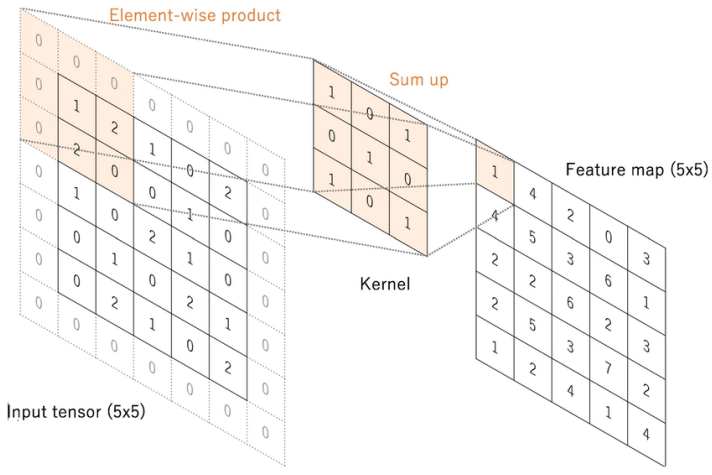


Figure 2.9: convolution operation

Where:

- n: size of the input $(n \times n)$
- f: size of the filter $(f \times f)$

- p: padding value

2. **Stride :** Stride denotes the pixel count that dictates the filter's movement across the input data in both horizontal and vertical directions. The stride regulates the manner in which the kernel convolves over the input. When the stride is one, the kernel is advanced over the input by a single element at each iteration. The stride is generally configured to ensure that the output volume is an integer.

- **Pooling Layer :** The primary function of the pooling layer is to diminish geographical dimensions and decrease the computational complexity of the network. Prevalent pooling methods encompass max pooling, which selects the maximum value from each filter, and average pooling, which employs the average operator. For an illustration of these pooling techniques, refer to Figure 2.10 .

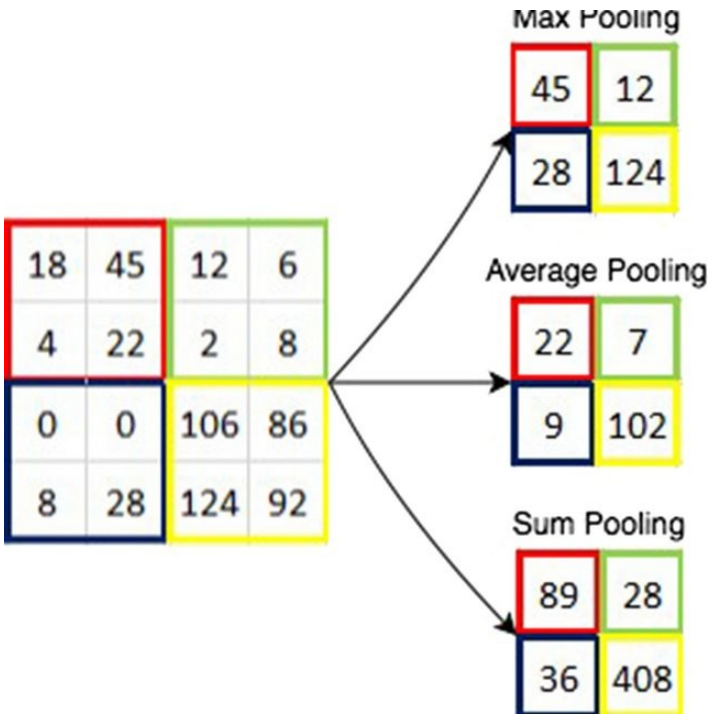


Figure 2.10: Illustration of Pooling Techniques

- **Fully connected layer:** The output is converted into a 1-dimensional array by the fully linked layer after feature extraction. Every element of the array is linked to every neuron in this layer. As a result, the network may learn to recognize trends in the data and provide accurate forecasts:

$$y = \sigma(W^T x + b) \tag{1.10}$$

where:

- σ: activation function
- W: weight matrix

– x: flattened input features

– b: bias vector

2.4.2 Long Short Term Memory

LSTM stands for Long Short-Term Memory, and it is a type of recurrent neural network (RNN) architecture that is commonly used in natural language processing, speech recognition, and other sequence modeling tasks. [3]
is a subset of RNNs developed specifically to deal with the vanishing gradient issue that plagues more conventional RNNs. Hochreiter and Schmidhuber's LSTM eliminates the drawbacks of conventional RNNs and old-fashioned ML techniques. The Keras package makes it easy to implement in Python. LSTMs control the flow of data via three critical gates:

- **F**orget Gate: Determines which past information should be discarded. For instance, if a scene ends, the model removes all associated data.
- **I**nput Gate: Controls which new information should be stored. When a new scene appears, this gate selectively retains critical details.
- **O**utput Gate: Combines long-term and short-term memory to generate predictions for the next hidden state or output. The overall architecture of the LSTM model is illustrated in the figure below.

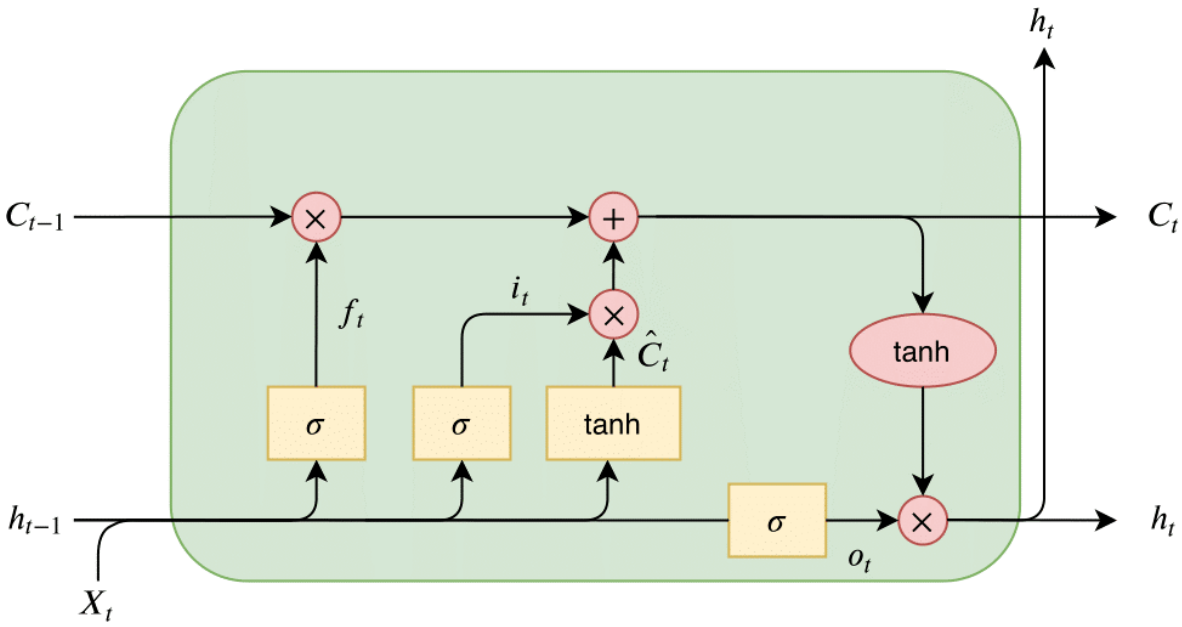


Figure 2.11: LSTM basic unit

2.5 Ensemble learning

Ensemble learning is a machine learning technique that aggregates two or more learners (e.g. regression models, neural networks) in order to produce better predictions. [29] Sometimes called committee-based learning, integrates multiple individual models to generate predictions with higher accuracy than those from a single model. This approach is grounded in the idea that a group of models (or "learners") collectively achieves superior performance compared to any individual model. Studies confirm its effectiveness across diverse machine learning frameworks, such as convolutional neural networks (CNNs), underscoring its versatility in improving predictive outcomes.

2.5.1 Types of Ensemble Models

There are two main schools of thought when it comes to machine learning ensemble learning approaches, according to the literature: **parallel** and **sequential**.

- **Parallel methods** Train each base learner independently of the others. These approaches develop basic learners concurrently, without interaction among the models.
- **Sequential methods** Sequentially train base learners, with each model endeavoring to rectify the flaws of its predecessor. Models are constructed incrementally for improved precision.

Parallel techniques are categorized into homogeneous and heterogeneous approaches. Homogeneous parallel ensembles use a same base learning approach to generate all component base learners. Heterogeneous parallel ensembles use diverse techniques to generate base learners.

Electoral Participation In what manner do ensemble approaches integrate base learners to form a final learner? Certain strategies, such as stacking, use distinct machine learning algorithms to train an ensemble learner from the foundational learners. A prevalent technique for aggregating base learner predictions is voting, specifically majority voting.

2.5.2 Ensemble learning techniques

- **Bagging :**is a homogeneous parallel technique sometimes referred to as bootstrap aggregating. It employs altered replicas of a specified training dataset to instruct multiple base learners using the same training procedure.12 The ensemble module of Scikit-learn in Python has methods for doing bagging, including BaggingClassifier.
- **Stacking:** Stacking, or layered generalization, is a heterogeneous parallel technique that exhibits meta-learning. Meta-learning involves training a meta-learner using the outputs of several base learners. Stacking particularly trains several base learners on the same dataset, using distinct training algorithms for each

learner. Every base learner generates predictions on an unobserved dataset. The initial model predictions are then aggregated and used to train a final model, referred to as the meta-model.

- **Boosting :** algorithms are a sequential ensemble method. Boosting has many variations, but they all follow the same general procedure. Boosting trains a learner on some initial dataset, d. The resultant learner is typically weak, misclassifying many samples in the dataset. Much like bagging, boosting then samples instances from the initial dataset to create a new dataset (d2). Unlike bagging, however, boosting prioritizes misclassified data instances from the first model or learner. A new learner is trained on this new dataset d2. Then a third dataset (d3) is then compiled from d1 and d2, prioritizes the second learner's misclassified samples and instances in which d1 and d2 disagree. The process repeats n times to produce n learners. Boosting then combines and weights the all the learners together to produce final predictions.

2.6 Conclusion

Within the scope of this chapter, the investigation of Artificial Intelligence (AI) and its subset, Machine Learning (ML), highlights the transformational significance that these technologies play in contemporary technology. Machine learning algorithms are able to adapt to a wide variety of data settings by using Supervised, Unsupervised, Semi-supervised, and Reinforcement Learning. This allows them to do jobs ranging from pattern recognition to autonomous decision-making among other things. The advancements in Deep Learning have further transformed artificial intelligence by using neural networks to interpret complicated and unstructured data. Ensemble Learning approaches, on the other hand, boost predicted accuracy by merging numerous models. The combination of these approaches exemplifies the adaptability and scalability of artificial intelligence systems. The combination of these professions promises to stimulate innovation across sectors, ranging from healthcare to autonomous systems, and will shape a future in which intelligent solutions meet ever-growing global concerns. This will be accomplished as these domains continue to advance.

Chapter 3

Stress Overview

3.1 Introduction

Stress is the physiological response of the human body characterized by significant anxiety or pressure in response to a hard situation. The clinical definition of stress refers to a psycho-physiological condition characterized by significant discomfort and anguish, which may lead to acute mental health issues such as depression or anxiety attacks.

A **stressor** is an event or condition surrounding an individual that may provoke stress. The effects of stress on an individual can be both beneficial and detrimental, contingent upon the management of stressful situations. Consequently, a scenario that is profoundly stressful for one person may elicit only a mild response in another. Furthermore, previous stressful experiences can function as a protective mechanism in similar future circumstances. For individuals who thrive in challenging environments, stress serves as a *adrenaline booster* and is frequently perceived as a positive response.

Stress that has a positive impact is called **Eustress**. It is a type of stress experienced when some stimulating incident is anticipated in the surrounding environment. It is marked by an increase in pulse rate but without any underlying sense of threat or fear. This form of stress typically occurs when the outcome is expected to be positive, such as during childbirth or competing for a promotion. Eustress can enhance mental abilities, motivate individuals to achieve their goals, and improve task efficiency. It pushes people out of their comfort zones, inspiring learning, personal growth, and resilience.

Eustress is a term that refers to stress that has a beneficial effect on the situation. The anticipation of an event that is expected to be stimulating in the environment around one is a kind of stress that might be experienced. An elevated heart rate is the defining characteristic of this condition, yet there is no underlying feeling of danger or anxiety. It is common for this kind of stress to manifest itself in situations when the result is anticipated to be favorable, such as when one is giving birth or vying for a promotion promotion. Eustress has the potential to strengthen mental capabilities, inspire people to attain their objectives, and increase the efficiency with which tasks are completed. It forces individuals to go beyond of their comfort zones, which provides an opportunity for learning, personal development, and resiliency.

Both external environments (e.g., the workplace) and internal stimuli (e.g., habitual behavior or emotions) can cause distress. Common sources include fear, anxiety about future events, recurring negative thoughts, unrealistic expectations, over-scheduling, job insecurity, and perfectionism. Personal issues such as illness, financial trouble, unemployment, or the loss of a loved one may also trigger distress. Early detection of duress is critical as prolonged stress can significantly impact an individual's quality of life.

Stress is also known as the **"fight-or-flight"** response, a survival mechanism enabling people to respond rapidly to threats or challenges. When confronted with a stressor, the body activates the sympathetic nervous system, releasing chemicals such as cortisol, adrenaline, and noradrenaline[36] . This biological response boosts strength, stamina, and reaction speed, allowing the person to determine whether to fight the danger or run from it. Hormones such as adrenaline increase heart rate, tense muscles, raise blood pressure, create shortness of breath, and heighten the senses.

3.2 Signs and Symptoms of Stress Overload

Over time, individuals may swiftly adapt to stress as it becomes more tolerable and susceptible to mitigation. They not only neglect to acknowledge the detrimental effects of prolonged exposure to even low stress levels, but they also overlook the damage this exposure inflicts on their health. Consequently, it is important to recognize the common warning signs and symptoms linked with stress overload. Table 3.1 presents a catalog of the indications and symptoms associated with stress overload. An overwhelming level of stress may induce severe depression in those who are susceptible to it. Generally, the existence of chronic stressful life situations may increase the likelihood of developing depression. The diagnosis of depression is facilitated by a number of different machine learning approaches, each of which plays a significant role.[26]-[27]

Table 3.1: Signs and Symptoms of Stress Overload

Cognitive Symptoms	Constantly being concerned, Anxious thoughts, forgetfulness, disorganization, and an inability to concentrate are all symptoms of anxiety. Challenges with money, Being pessimistic, or focusing only on the negative aspects of a situation, Having trouble concentrating, Having poor judgment.
Physical Symptoms	Discomforts and aches, Feeling sick and lightheaded, The common cold and influenza, Constipation or diarrhea, or both Chest discomfort, and Rapid pace of heartbeat, Distention of the bowels, a sensation of choking, muscles that are rigid or tight, Crushing one's teeth, Urinating often and with a sense of urgency, A feeling of exhaustion, The loss or increase of weight.
Affective Symptoms	Sensations of tension, irritability, or anger, as well as Feelings of restlessness, anxiety, and an inability to relax anxiety and agitation, depression, general sadness, and general unhappiness feelings of irritability, loneliness, and isolation, Additional issues pertaining to mental and emotional health.
Behavioral Symptoms	Issues with sleeping, Uncertainty about the completion of labor responsibilities, Abnormalities in appetite, such as not eating at all or eating an excessive amount Putting off the completion of tasks and avoiding obligations, An increase in the use of alcoholic beverages, illicit substances, or smokes Participating in more anxious habits, such as biting one's nails, fidgeting, and pacing about the room.

3.3 Biosignals related to stress

A state of imbalance between the sympathetic and parasympathetic nervous systems may occur as a consequence of the stimulation of the autonomic nervous system that occurs under stressful situations. As a consequence of this, the symptoms of stress often consist of physiological and physical reactions that originate from the nervous system. Physiological biosignals are able to measure the deformation of the body that is caused by muscle movement. Included among them are:

- Pupil size,
- Eye movements or blinks,
- Respiration,
- Facial expressions
- Body and extremity semivoluntary position/movements.

Physiological signals are more directly related with body vital functions, such as :

- Electrocardiogram (ECG),
- Electroencephalogram (EEG),
- Electrodermal Activity (EDA),
- Respiration and Breath Rate (BR),
- Photoplethysmography (PPG).

The following subsections classify biosignals based on their anatomical origin, namely those obtained from the brain, the heart, and other bodily regions. Figure 3.1

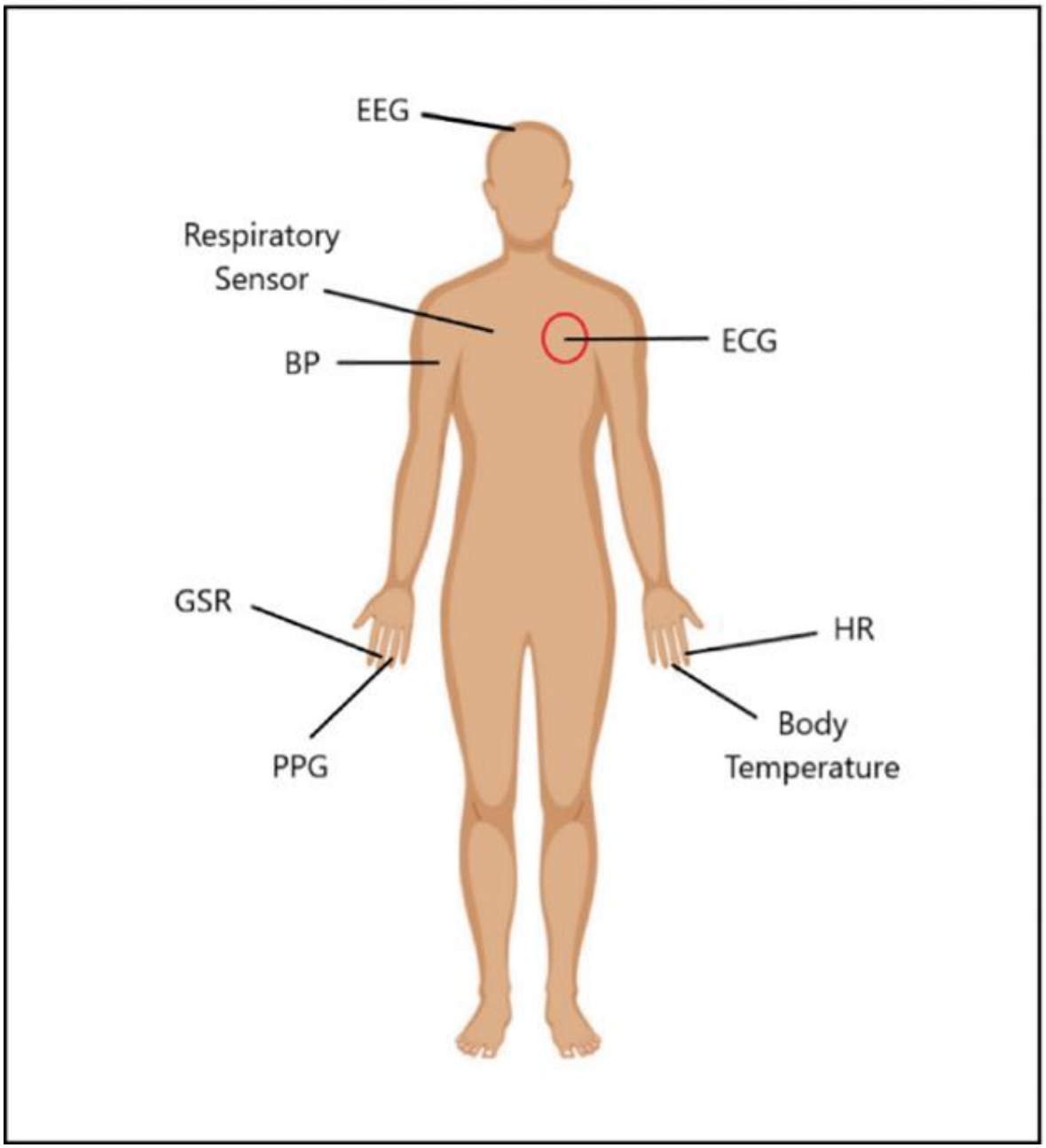


Figure 3.1: Schematic schematic illustrating typical locations of wearable sensors on the human body

3.3.1 Heart Rate (HR)

The electrocardiogram (ECG) is a diagnostic examination that captures the heart's electrical activity. This is a non-invasive treatment that entails the placement of tiny electrodes on the patient's chest, arms, and legs, which detect the Electrical impulses produced by the heart's cells during contraction and relaxation are sent to a device that records the information. An ECG is a time-series representation of the heart's electrical activity, with each pulse shown as a sequence of electrical waves. The variations

in electrical potential difference (voltage) during the depolarization and repolarization of cardiac fibers are captured by electrodes placed on the chest surface; this voltage constitutes the amplitude of the ECG signal. The ECG, seen in Figure 3.2, has three fundamental waves: P, QRS, and T. These waves represent the distant field generated by certain electrical activities on the heart surface, namely, atrial depolarization (P), ventricular depolarization (QRS complex), and ventricular repolarization (T).
Differences in the speed of wavefront propagation through the cardiac cycle are reflected by different frequencies content of ECG waves. The content of T wave lays mostly within a range from zero (DC) to 10 Hz. The content of P wave is characterized by 5-30 Hz frequencies. The content of QRS usually contains within 8-50 Hz frequencies while abnormal ventricular conduction is characterized by high frequencies (above 70Hz), forming notches on the QRS [40]. Stress leads to the activation of the Sympathetic Nervous System (SNS), resulting in the increase of heart rate and its force of contraction. As a result, the amount of blood circulates faster through the body to deliver more oxygen to the organs and skeletal muscles as an attempt to eliminate the stressor [17]. During times of stress, the body releases hormones like adrenaline and cortisol, which can increase heart rate, blood pressure, and cardiac output. Alterations in the body's physiology may be identified by an ECG, which will reveal an elevation in heart rate, changes in the morphology of the waves on the ECG tracing, and more irregularities. Heart rate (HR) is the most prevalent

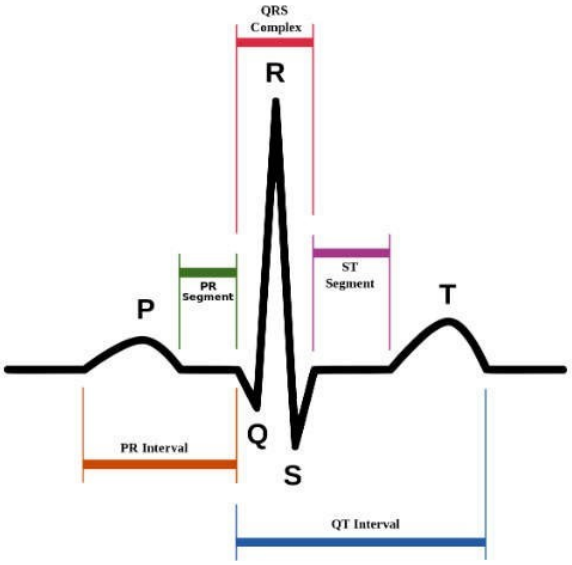


Figure 3.2: Typical ECG signal with characteristics peaks P,Q,R,S,T
[41]

A direct and effective method to assess stress levels is via measuring heart rate, which dramatically elevates under stressful conditions. Heart rate (HR) is among the most extensively used indicators of human stress documented in the literature, as seen in section **??** . Heart rate is defined as the quantity of heartbeats per minute (bpm). The RR interval of the ECG signal, defined as the duration between successive heartbeats, has an inverse correlation with an individual's heart rate. An ECG is an effective instrument for assessing the effects of stress on the body and identifying

possible cardiac issues associated with stress. Heart Rate is the most widely adopted and straightforward measure to estimate levels of stress. Alternatively, the mean RR interval can be used, having an inverse relationship with heart rate. According to [17], the heart rate is the most important feature used to detect stress, being a reliable measure for the arousal part of stress.

Heart Rate is the most widely adopted and straightforward measure to estimate levels of stress. Alternatively, the mean RR interval can be used, having an inverse relationship with heart rate. According to [17], the heart rate is the most important feature used to detect stress being a reliable measure for the arousal part of stress. The HRV, instead, has a chaotic behavior in states of anger, anxiety or sadness whose rhythmicity can be described by a measure known as cardiac coherence. SDNN and RMSSD are reduced during stress conditions.

In frequency domain, instead, the LF band is modulated by both sympathetic and parasympathetic activity, while the HF band corresponds only to parasympathetic activity. Thus, the ratio LF/HF is considered a distinctive approach for the sympathetic modulation and is the more prominent feature in frequency domain increasing during stress condition.

Table 3.2: Main Features Related to stress detection for ECG signal

Features	signification
HR	Heart Rate
HRV	Heart Rate Variability
LF	Low Frequency
HF	High Frequency
SDNN	Standard Deviation of Normal-to-Normal Intervals

3.3.2 Electroencephalogram (EEG)

The electroencephalogram (EEG) records the brain's electrical activity via the scalp. This is a non-invasive treatment that involves the application of tiny electrodes to the patient's scalp. The electrodes sense the electrical impulses. created by the brain's neurons during intercommunication and sent to a computer that records the data. The documented data is shown visually as a sequence of waves that correlate to various levels of cerebral activity. These waveforms reflect the cortical electrical activity, which is a widely used technique to estimate changes in neurophysiological activity associated with external stimuli and/or with the performance of specific tasks [17] and has a typical shape as shown in Figure 3.3. An EEG exhibits a frequency range of 0.01 to approximately 100 Hz and fluctuates between a few microvolts and roughly 100 µV. The slow components near 0.01 Hz correspond to slow cortical potentials, which are typically excluded in clinical practice; however, they may hold significance for brain-computer interfaces. The predominant way for classifying EEG waveforms is based on frequency, with the most extensively examined waveforms being [2]:

- Delta waves, with a frequency range of 0.5 to 4 Hz, are physiologically seen during deep sleep and are particularly pronounced in the frontocentral areas of the skull.

- Theta waves, with a frequency range of 4 to 7 Hz, are associated with tiredness and the first phases of sleep, mostly seen in the fronto-central regions of the skull.
- The alpha rhythm, with a frequency range of 8 to 12 Hz, is predominantly seen in the occipital area during relaxation or states of little cognitive load and emotional stress. The characteristic element of the typical background rhythm in adult EEG recordings;
- Sigma activity, characterized by a frequency range of 12 to 16 Hz, is often referred to as "sleep spindles." This activity may be categorized as slow (12 to 14 Hz) or rapid (14 to 16 Hz) and is most frequently seen in the fronto-central areas of the scalp.
- The beta rhythm, characterized by a frequency range of 13 to 30 Hz, is the predominant neural oscillation seen in healthy adults and children during situations of heightened cognitive processing or alertness. It is most pronounced in the frontal and middle parts of the scalp and diminishes in amplitude towards the posterior areas.

Stress situations are seen as diminishing alpha activity and augmenting Beta wave activity. Certain variables, such as sleep deprivation and the execution of a cognitive activity, may only influence the lower alpha power. The lower alpha power (8-10.5 Hz) is present, but the greater alpha power (11-12 Hz) may be diminished in both circumstances. The influence of elevated frequency rhythms (ranging from 30 to 70 Hz) may serve as a precise indicator of the amount of stress reaction [17].
Alpha band activity predominates during relaxation when cognitive demands are low, but beta band activity is substantial in conditions of severe strain or alertness. Stress is connected with beta waves in the temporal region of the brain. In instances of negative mood or depression, alpha and beta band activity predominates.

Extended exposure to stress may adversely affect brain function and contribute to the emergence of many neurological disorders, including anxiety, depression, and post-traumatic stress disorder (PTSD). These diseases may also be identified and diagnosed by alterations in the electrical patterns captured by an EEG. The EEG offers comprehensive and dependable insights into the determinants of stress, as the continuous real-time recording of EEG signals significantly enhances the comprehension of stress patterns in field subjects, given that these signals swiftly react to various stressors. Consequently, EEG may be very beneficial in examining environmental stresses or creating online physiological monitoring systems.

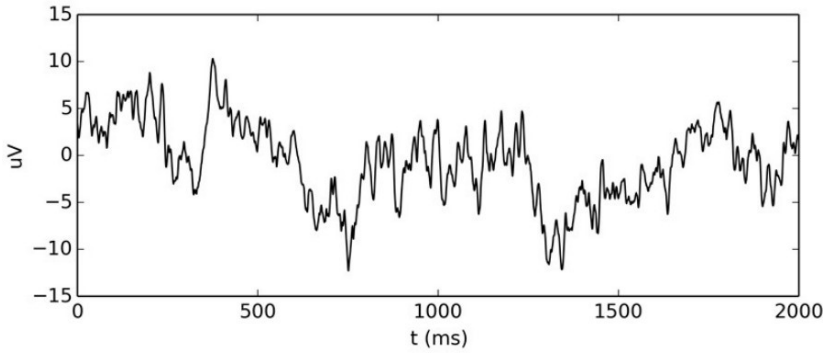


Figure 3.3: Typical raw EEG acquired with biosignalplux [8]

3.3.3 Electrodermal Activity (EDA)

Sweat glands secrete moisture via pores to the skin's surface when activated. When the equilibrium of positive and negative ions in this produced fluid changes, the electrical current conducts more efficiently. This leads to decreased skin resistance, or conversely, heightened skin conductance. Electrodermal activity (EDA), often referred to as galvanic skin response (GSR), is a physiological assessment of electrical conductivity in the skin. An EDA sensor quantifies variations in skin conductance, which are induced by fluctuations in sweat secretion and sweat gland activity due to changes in sympathetic nervous system activity. Even minimal, imperceptible sweating can affect skin electrical conductivity[17]. Electrodermal activity (EDA) is quantified using electrodes affixed to the skin, usually on the fingers or palms, and is often used in both research and clinical environments as a physiological indicator of emotional arousal and stress, as seen in Figure 3.5. EDA has been used to investigate several phenomena, including emotional reactions, cognitive processing, and physiological sensitivity to stress. The EDA signal measurements consist of the convolution of two signals, as seen in Figure 3.4:

- Skin conductance level (SCL), tonic component: this signal pertains to the slower-acting elements and baseline properties of the signal;
- The phasic component of Skin Conductance Response (SCR) pertains to the rapidly fluctuating aspects of a signal.

The SCL signal determines the fundamental level of the signal. Alterations in the SCL are believed to indicate overall modifications in autonomic arousal, a physiological reaction that transpires when the autonomic nerve system (ANS) is stimulated, leading to variations in body processes that are involuntary. The tonic level of the EDA signal might vary considerably across people. The tonic level alone is not really informative.

The SCR is intricately linked to the function of the sweat motor system, which is concurrently coupled with the parasympathetic nervous system.

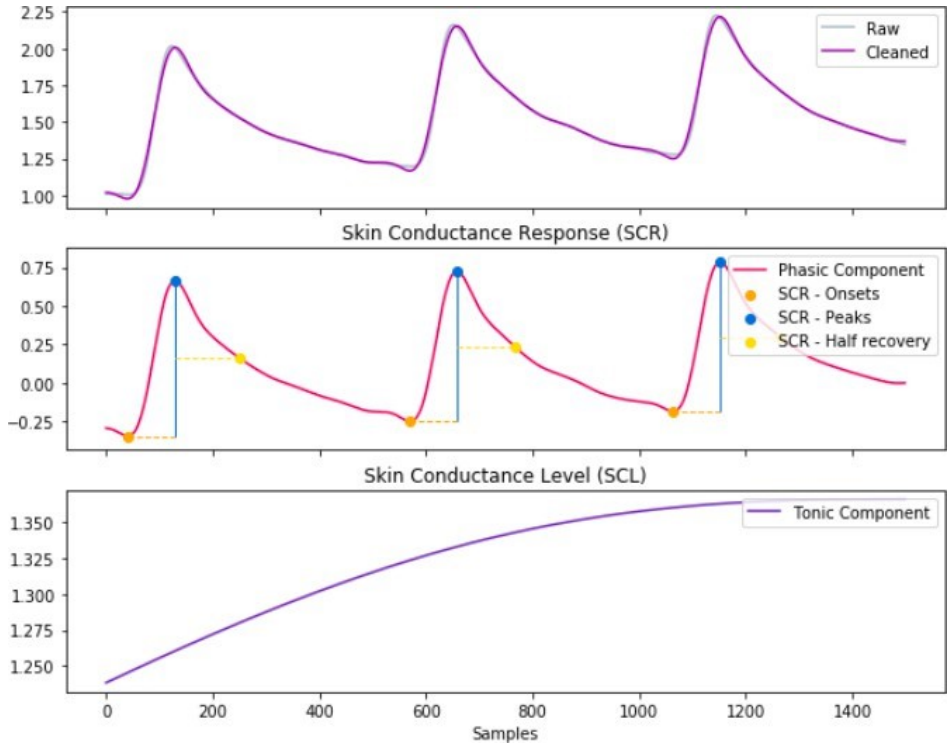


Figure 3.4: Typical components of EDA signal.

[5]. The phasic component is responsive to certain emotionally stimulating events; these bursts may occur 1-5 seconds after the initiation of emotional stimuli. The skin response data exhibits a different pattern based on the individual's condition is regarded as a dependable approach for measuring stress. The SCR component of EDA increases when faced with an emotionally stimulating scenario.

A robust association exists between electrodermal activity and stress. When an individual experiences stress, their sympathetic nervous system activates, resulting in increased perspiration on the skin.leading to an elevation in the SCL and SCR components of electrodermal activity.

When an individual experiences stress, the two most significant characteristics of the EDA are both tonic components.The SCL and the phasic component of SCR increase owing to heightened skin moisture levels. As per [17] , SCL was deemed the most efficacious stress correlate among the characteristics derived from HRV, RSP, and EMG. Additional prevalent characteristics used in stress research include SCR frequency, SCR amplitude, SCR latency, SCR rising time, SCR half recovery, and SCR reaction initiation. Researchers have discovered that individuals experiencing elevated chronic stress, such as those with post-traumatic stress disorder (PTSD), have greater levels of electrodermal activity (EDA) compared to those without such conditions. Furthermore, EDA has been used in biofeedback therapy to assist persons in managing stress and anxiety. By monitoring their EDA, people may discern and modulate their physiological reactions to stress, so enhancing their overall well-being.

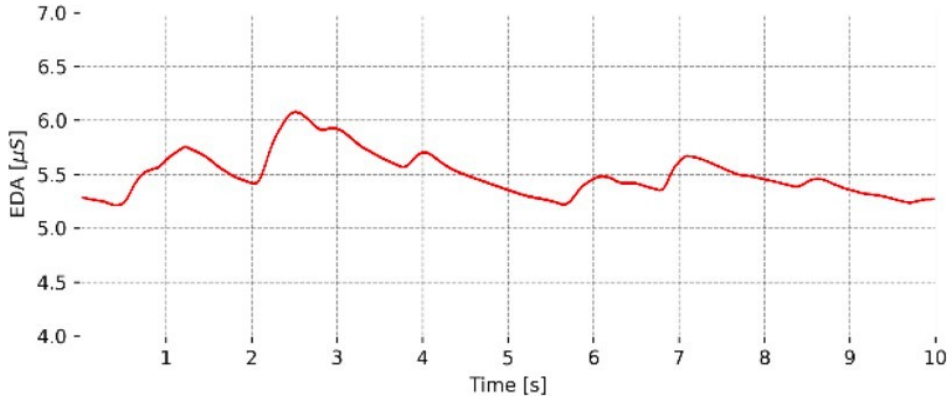


Figure 3.5: Typical raw EDA acquired with biosignalplux [7]

While thermoregulation is a primary function of sweating, it is also elicited by stimuli, such as emotionally charged imagery. This phenomenon is referred to as emotional perspiration. Perspiration . The secretion, indicative of arousal fluctuations, is automatically regulated by the autonomic nervous system (ANS) to fulfill behavioral demands. During periods of stress, both the tonic component of skin conductance level (SCL) and the phasic component of skin conductance response (SCR) elevate owing to an increase in skin wetness. The mere anticipation of an unpleasant or stressful event might provoke elevations in the EDA. The maxima of SCR typically manifest between 1.5 and 6.5 seconds after the initiation of stressor stimuli. [17] The human skin's reaction is not consciously controlled and is influenced by changes in the subject's sweating pattern, hence indicating the activity of the sympathetic nervous system. GSR measuring sites are areas of the body abundant in sweat glands. Various places on the human body are available for measuring GSR.

3.3.4 Respiration and Breath Rate (BR)

The respiration signal denotes the physiological process of breathing, including the inhalation of oxygen-rich air and the exhalation of carbon dioxide-rich air. Respiration is regulated by the respiratory center in the brainstem, which gets information from numerous sensors in the body that monitor fluctuations in oxygen levels. Carbon dioxide concentrations in the blood, together with other variables such as pH and temperature. This signal often exhibits the form seen in Figure 3.6. The respiratory rate is the frequency of breathing, often quantified by physically tallying chest wall movements per minute. The respiratory center is situated in the medulla oblongata (lower brainstem) and regulates breathing on a minute-to-minute basis [23]. This rate may fluctuate based on many variables, including age, physical activity, and general health. The standard respiratory rate for resting adults is generally between 12 and 20 breaths per minute, however for youngsters, it ranges from 20 to 30 breaths per minute.

During stress, respiration rate typically escalates with emotional arousal and diminishes with relaxation, but stressful circumstances may induce transient disruptions in breathing. Negative emotions, including stress, are associated with abnormalities

in respiratory patterns, an increase in minute volume, a transition from abdominal to thoracic breathing, and accelerated, shallower respiration [17]. The oxygen consumption rate may be derived from an individual's respiratory rate and is regarded as a valid indicator of human stress, since oxygen demand escalates under stressful conditions [4].

The Respiration Signal (RESP) primarily monitors breath rate and breath depth (amplitude). Under stress, the respiratory rate typically rises with emotional arousal and decreases with relaxation, whereas stressful circumstances may cause transient interruptions in breathing. The breath rate (BR) exhibits considerable alterations under stress, as seen in 1.4. Moreover, unpleasant emotions such as stress are associated with abnormalities in respiratory patterns, an increase in minute volume, a transition from abdominal to thoracic breathing, and accelerated, shallower respiration [17].

In stressful situations, the body's "fight or flight" reaction is triggered, resulting in an elevated respiratory rate. The elevated respiration rate may be seen by alterations in the respiration signal, including an increase in both the number and depth of breaths. These modifications aim to provide the body with more oxygen required to react to the stressful circumstance. Prolonged exposure to stress may adversely affect respiratory function and contribute to the onset of numerous respiratory diseases, including asthma, COPD, and hyperventilation syndrome. These disorders may also be identified and diagnosed by alterations in the respiratory signal.

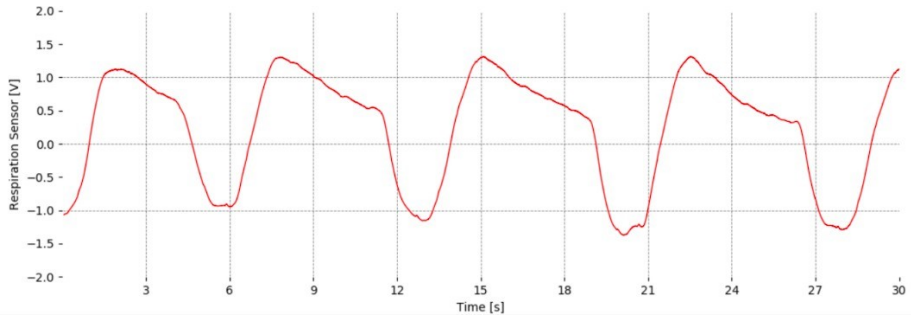


Figure 3.6: Raw Respiration signal acquired with biosignalplux [9]

Extracting respiratory signals from the electrocardiogram (ECG) or photoplethysmogram (PPG) using either feature- or filter-based methods allows one to estimate breathing rate (BR). The three idealised kinds of respiratory modulation of the ECG and PPG are baseline wander (BW), amplitude modulation (AM), and frequency modulation (FM). Should the amplitude of the respiratory signal be too low in relation to the underlying noise, the signal could not be identifiable from the noise, hence hindering the exact calculation of BR. Respiration alters the morphology of the ECG by changing the electrical impedance of the chest caused by filling and emptying of the lungs.

Respiration alters the morphology of the ECG by changing the electrical impedance of the chest as the lungs fill and empty. Monitoring of breath control has also been suggested as a means to reduce the HRV oscillations caused by respiration; respiratory rate may also be utilized as a characteristic; combine it with other biosignals to determine the stress level of the patients.

3.3.5 Photoplethysmography (PPG)

Photoplethysmography (PPG) is an optical non-invasive method measuring variations of skin hue associated with concurrent changes in blood volume in subcutaneous blood vessels during the cardiac cycle. This measurement provides valuable information about the cardiovascular system [17] .

A typical PPG device has a photodetector that receives the reflected light from a light source generating pulses. The light source sends light to a tissue; the photodetector then records the reflected light from the tissue. Reflected light varies in proportion to blood volume changes[11] . The PPG signals allow for the extraction of pulse rate (PR), pulse rate variability (PRV), and blood pressure (BP); they are also used for HRV parameter assessment because they show great temporal peak correlation with respect to ECG.

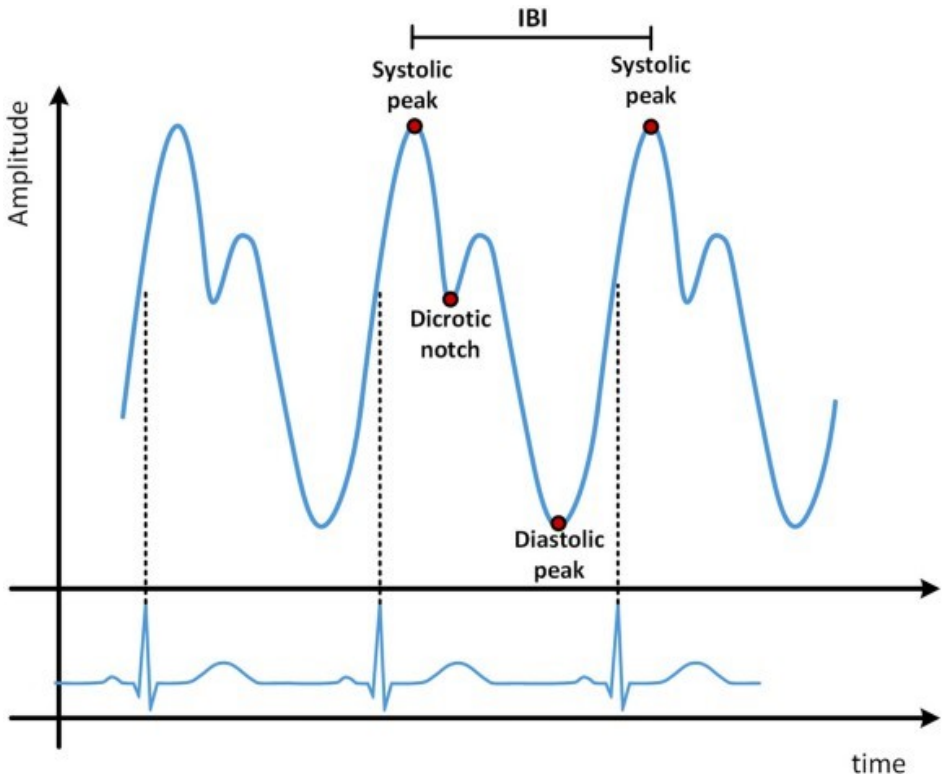


Figure 3.7: Raw PPG signal and its corresponding ECG.

Basically, PPG signal is made up of four components: diastolic points, systolic points, dicrotic notch, and dicrotic wave as seen in Figure 3.7. The diastolic and systolic points are the key mechanisms in PPG signal as they may provide valuable insights on the cardiovascular system. Beginning with a valley and ending with the pulse wave systolic peaks, the systolic phase runs until the pulse wave end, which is signaled by a second valley at the conclusion of the diastolic.

phase. While the time duration between two consecutive systolic points in the signal determine the instantaneous heart rate of an individual [20]. The stress-induced vascular response index (sVRI), which is a PPG-based measure is proposed to assess stress levels. Stress can also be reflected in peripheral vasoconstriction, being related

to decreased pulse wave amplitude (PWA) in PPG signals [20].

3.4 Related Work

The COVID-19 pandemic has heightened interest in telehealth, leading to research using cameras for the estimation of rPPG signals and real-time monitoring, addressing the need for non-invasive, contactless, and accessible methods for stress assessment [34][24]. rPPG provides a non-invasive method for remotely assessing BVP. This method needs only a camera and a source of ambient light. With this, HRV measures, pulse rate,and breathing rate can be measured using an everyday camera for facial video analysis to remotely detect and monitor stress. [[13] -[28]- [42]] There have been a growing number of research papers. For example, Benezeth et al. [6] proposed an rPPG-based algorithm that estimates HRV using a simple camera, showing a strong correlation between the HRV features and different emotional states. Similarly, Sabour et al. [30] proposed an rPPG-based stress estimation system with an accuracy of 85.48%. Some other works on the use of rPPG are encouraging, indicating that non-contact measures of some human physiological parameters (e.g., breathing rate (BR) and Heart Rate (HR)) are promising and have great potential for various applications, such as health monitoring [[21]–[1]] and affective computing [[43]–[25]].Although these efforts are commendable, this substantially propels The domain is advanced by the introduction of Hybrid Deep Learning (DL) networks and models for rPPG signal reconstruction and Heart Rate (HR) estimation. This innovative method demonstrates a significant improvement in precision and efficacy in stress detection, with an accuracy of up to 95.83%. use the UBFC-Phys dataset. The use of Hybrid DL networks signifies a contribution, providing improved functionalities for signal reconstruction and stress categorization. Given these factors, rPPG is well suited for both commercial and quotidian applications, with the notable benefit of monitoring ECG and photoplethysmography (PPG).

Traditional methods of stress detection include limitations that are incompatible with contemporary lives and the need for real-time monitoring. These approaches are intrusive, susceptible to bias, entail significant expenditures, and need time-consuming travel to clinical environments. In the last twenty years, there has been a significant transition towards technology-based methods for more efficient, cost-effective, and less invasive stress monitoring that align with contemporary lifestyles. Wearable gadgets, smartphone apps, and Machine Learning (ML) algorithms have transformed the detection and quantification of stress. One approach is measuring HRVusing wearable devices such as smartwatches, fitness trackers, and chest straps, allowing continuous and long-term monitoring of stress levels [38] [39] [10] [32] [18] . Typically, as HRV measures are inherently nonlinear, ML algorithms and other statistical data-driven methods such as Modified Varying Index Coefficient Autoregression Model (MVICAR) [12] can be applied in stress detection systems. ML algorithms have enabled accurate and efficient HRV-based stress detection and classification systems [31] [35] [22] . EDA, which measures the electrical activity of sweat glands, is another method that can be monitored with wearable devices, providing continuous and real-time monitoring of stress levels. Mobile applications using EDA-based biofeedback help individuals manage stress by providing real-time feedback and stress reduction techniques [38] [14] .Nonetheless, EDA assessment is influenced by ambient variables, dermal conditions,

and pharmacological agents, hence impacting its accuracy.

Wearable and contactless technologies are viable options for stress assessment, offering simple and non-invasive means for continuous monitoring. Nonetheless, the quality and precision of the data produced by various devices may differ. A significant constraint to adaptation rPPG is characterized by a reduction in the signal-to-noise ratio, necessitating sophisticated signal processing techniques. Numerous papers lack peer review and validation in clinical environments, prompting questions over the credibility of the findings. Wearable devices may be affected by variables such as motion, temperature, and perspiration, resulting in imprecise data; also, their use during sleep or strenuous activity presents a significant constraint. Individuals with skin sensitivities, allergies, or certain health issues may also find the use of these devices unbearable.

3.5 Conclusion

Stress is a multifaceted physiological and psychological reaction that profoundly affects individual well-being. This chapter offered an extensive examination of stress, encompassing its definitions, classifications, symptoms, and the physiological biosignals used for its identification, including heart rate (HR), electroencephalography (EEG), electrodermal activity (EDA), breathing rate (BR), and photoplethysmography (PPG). Biosignals are crucial for non-invasive, real-time stress evaluation utilizing contemporary wearable and remote sensor technology. We also examined pertinent research that underscores the progression of stress detection methodologies, transitioning from conventional clinical evaluations to AI-enhanced and camera-based strategies. This foundation facilitates the advancement of intelligent, real-time systems capable of precisely monitoring and managing stress in daily environments.

Chapter 4

Conception

4.1 Indroduction

This chapter will address the foundational framework of our Stress Detection System model and the strategies included inside it. We have meticulously created this framework and further divided it into numerous distinct sections, each carefully crafted to build upon its predecessor. This area encompasses data collecting, data preparation, model training, and real-world evaluation. It is important to execute each phase to assure the overall precision, adaptability, and reliability of the detecting system.

4.2 Suggested Approach

The Real-Time Stress Detection System project is advancing with the assistance of a meticulously developed framework. This paradigm governs the dataset, pre-processing, model training, and assessment of the real-world application. These stages are interconnected and must be executed in the specified sequence. Our inquiry was founded on a robust foundation provided by the extensive collection of annotated datasets acquired at the beginning of the procedure. Upon completing the data collection approach, we implemented several pre-processing steps to guarantee the data was adequately prepared for analysis. Subsequently, we entered the training phase, when we used sophisticated methodologies to adequately train our model for a real-time stress detection feature. Ultimately, we conducted thorough evaluations of the model's performance to ascertain its utility and reliability in real-world scenarios. We have created a schematic representation of our development framework's architecture to enhance our understanding of it.

Figure 4.1 illustrates this depiction for your perusal. The purpose of this visual aid is to offer an explanation of the three fundamental steps that serve as the foundation for the development and implementation of our system.

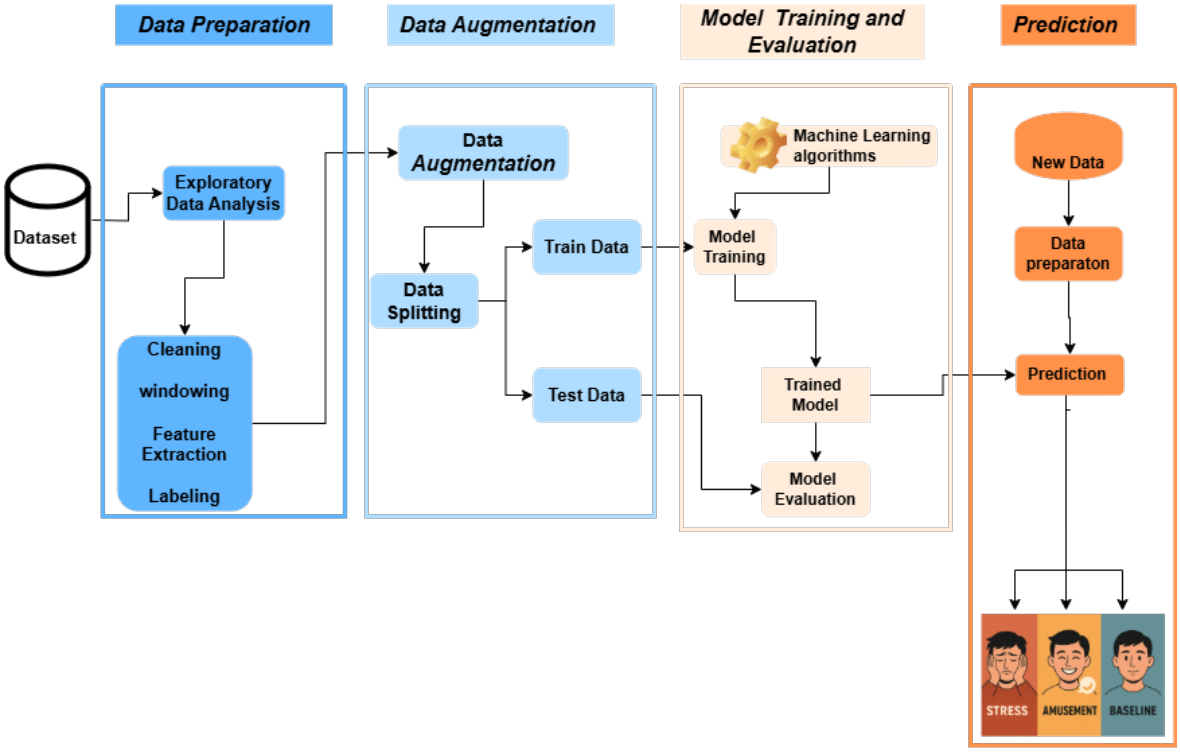


Figure 4.1: Schematic Representation Development of a Real-Time Stress Detection System.

4.3 Phases

4.3.1 Preprocessing

The photoplethysmographic (PPG) data is essential for identifying physiological changes associated with stress. A systematic preprocessing workflow was established to ready the raw photoplethysmography (PPG) data for feature extraction and ensuing machine learning analysis. This section provides a comprehensive review of the preprocessing methods.

a) **Signal Cleaning**

To enhance signal quality, the raw PPG signal is passed through a digital low-pass filter to suppress high-frequency noise and motion-related artifacts. A Butterworth filter is often used due to its smooth frequency response:

$$y(t) = \text{Butterworth}_{\text{low-pass}}(x(t), f_c) \tag{4.1}$$

where $x(t)$ is the original PPG signal, f_c is the cutoff frequency (typically 3–5 Hz), and $y(t)$ is the filtered signal.

b) **Segmentation via Sliding Window Approach**

To facilitate localized and consistent analysis of the PPG signal, a sliding window technique was employed. The signal was segmented into fixed-duration windows with overlapping intervals as follows:

- **Window duration:** $T = 30\,\text{seconds}$, resulting in $N = T \times 64 = 1920$ samples per window.
- **Window shift:** $\delta = 5\,\text{seconds}$, yielding a shift of 320 samples between consecutive windows.

The number of windows W that can be generated from a signal of length L samples is computed as:

$$W = \left\lfloor \frac{L - N}{\delta N} \right\rfloor + 1 \tag{4.2}$$

Illustrative Example: For a 10-minute BVP recording ($L = 64 \times 600 = 38400$ samples), the number of windows is:

$$W = \left\lfloor \frac{38400 - 1920}{320} \right\rfloor + 1 = 115 \tag{4.3}$$

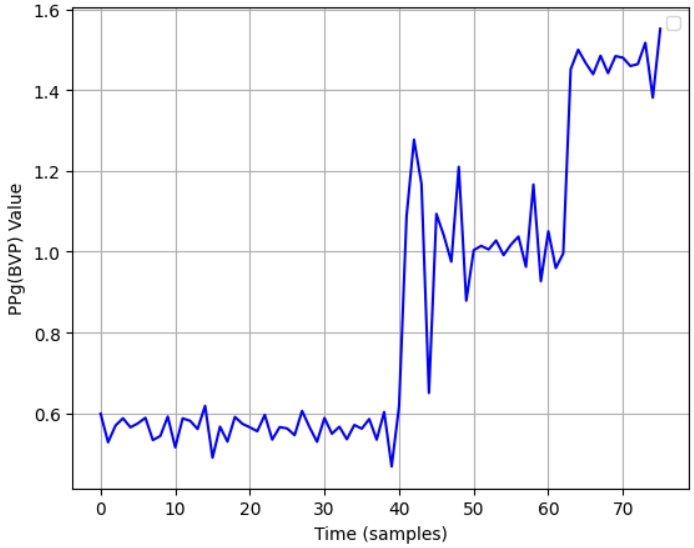


Figure 4.2: Original PPG (BVP)Signal (Before Windowing)

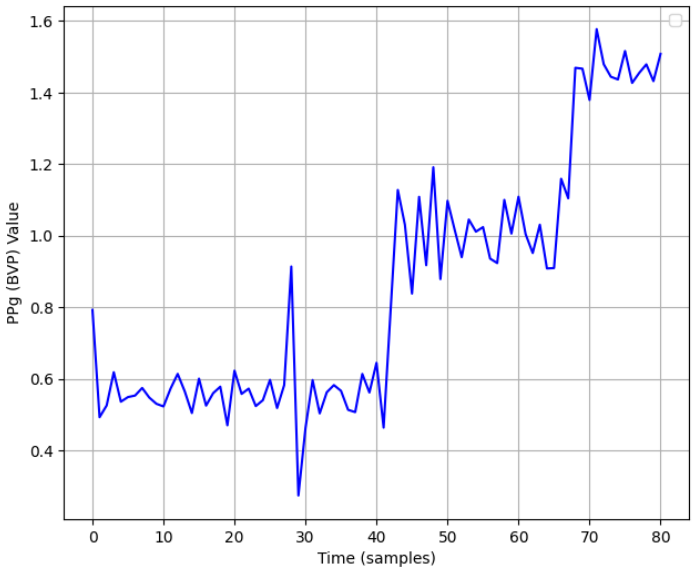


Figure 4.3: PPG(BVP) Signal with Sliding Windows (After Windowing)

c) **Statistical Feature Extraction**

From each 30-second window, we extracted a set of statistical features that characterize the signal dynamics:

- **Mean (μ):** The mean represents the average signal amplitude within a window. It provides a central tendency measure of the PPG waveform, which may shift due to changes in blood volume caused by stress or relaxation. It is computed as:

$$\mu = \frac{1}{N} \sum_{i=1}^{N} x_i \tag{4.4}$$

  where x_i is the i-th sample in the window and N is the total number of samples.

- **Standard Deviation (σ):** The standard deviationbh quantifies the dispersion or variability of the signal around its mean. Higher variability may indicate irregularities in pulse amplitude or increased physiological arousal. It is calculated using:

$$\sigma = \sqrt{\frac{1}{N-1} \sum_{i=1}^{N} (x_i - \mu)^2} \tag{4.5}$$

- **Minimum ($x_{\min}$):** The minimum value in the window represents the lowest BVP amplitude recorded, which can reflect brief drops in blood volume due to physiological events such as vasoconstriction.

- **Maximum ($x_{\max}$):** The maximum value indicates the peak PPG amplitude within the window, which can be associated with heartbeat strength or vessel dilation.

  Together, the min and max values define the amplitude range of the signal within a given window:

$$x_{\min} = \min(x), \quad x_{\max} = \max(x) \tag{4.6}$$

These features were chosen for their effectiveness in representing cardiovascular responses related to autonomic nervous system activity during stress.

d) **Label Assignment and Dataset Construction**

Each feature vector extracted from a PPG window was labeled according to the participant's state during that period:

- **Label 1:** Baseline
- **Label 2:** Stress
- **Label 0:** Amusement

The final dataset consisted of all labeled feature vectors across subjects, which was then used for training and evaluating stress classification models.

4.3.2 Data Augmentation

To enhance the generalization capability of the stress detection model and mitigate overfitting, we performed data augmentation on particular derived PPG signal features. The employed features included: `PPG_mean`, `PPG_std`, `PPG_min`, and `PPG_max`. These statistical descriptors were chosen for their physiological relevance and correlation with cardiovascular dynamics under stress. The augmentation pipeline included the following transformations:

Gaussian Noise Injection

We introduced small random variations to the PPG features by adding Gaussian noise, emulating sensor noise and physiological variability. For a given feature vector x, the noisy version x_{noisy} was computed as:

$$x_{\text{noisy}} = x + \mathcal{N}(0, \sigma^2) \tag{4.7}$$

Where:

- $\mathcal{N}(0, \sigma^2)$ is a zero-mean Gaussian distribution with standard deviation $\sigma = 0.01$.
- This technique artificially expands the dataset and simulates real-world inconsistencies in signal acquisition.

Temporal Stretching

To simulate changes in heart rate or physiological pacing, we scaled the PPG values by a constant factor. This is equivalent to a time-warping transformation in the temporal domain. For each sample:

$$x_{\text{stretch}} = \alpha \cdot x, \quad \text{where } \alpha = 1.1 \tag{4.8}$$

This operation stretches the magnitude of the original signal, mimicking prolonged cardiac cycles or slower pulse dynamics.

Temporal Shifting

A time shift was applied to simulate delays or phase shifts in signal acquisition. The shifting process preserved the original signal structure while altering its alignment over time. This was achieved via:

$$x_{\text{shift}}[t] = x[t + \tau], \quad \text{with } \tau = 5 \text{ samples} \tag{4.9}$$

Missing values introduced by shifting were filled with zeros to maintain consistent input dimensions.

Concatenation of Augmented Samples

For each original sample, three augmented versions were created: noisy, stretched, and shifted. These were then concatenated with the original feature vectors, resulting in a fourfold increase in training data size. Labels were preserved during augmentation using a majority vote across one-hot encoded columns (`'0'`, `'1'`, `'2'`).

Final Output

The final augmented dataset was saved in a consolidated CSV file with dimensions proportional to the original dataset multiplied by the number of augmentations. This enriched dataset was used in downstream model training to improve robustness and generalization.

4.3.3 LSTM Model for Stress Classification

In this study, a **Long Short-Term Memory (LSTM)** neural network was implemented to classify stress-related states based on Photoplethysmography (PPG) signal features. LSTM networks are a type of Recurrent Neural Network (RNN) capable of learning temporal dependencies in sequential data, which makes them particularly effective for time-series physiological signals.

Model Architecture

The architecture of the LSTM model used in this work is encapsulated in Table 4.1.

Table 4.1: LSTM Model Architecture

Layer (type)	**Output Shape**	**Parameters**	**Description**
lstm_1 (LSTM)	(**None**, 64)	16,896	LSTM layer with 64 units; processes input shape (4, 1) representing short PPG sequences
dense_6 (Dense)	(**None**, 64)	4,160	Fully connected layer with ReLU activation
dropout_3 (Dropout)	(**None**, 64)	0	Dropout layer with a rate of 0.5 to reduce overfitting
dense_7 (Dense)	(**None**, 3)	195	Output layer with softmax activation for classifying into 3 categories: baseline, stress, and amusement

- Total params: 63,755 (249.05 KB)
- Trainable params: 21,251 (83.01 KB)
- Non-trainable params: 0 (0.00 B)
- Optimizer params: 42,504 (166.04 KB)

The basic objective of the LSTM model is to use short-term PPG signal sequences to classify an individual's physiological state, including baseline, stress, or entertainment conditions. To get accurate classification, LSTM utilizes its temporal learning abilities, enabling the model to detect subtle time-dependent patterns.

LSTM-based attention model architecture

Table 4.2: Summary of the proposed LSTM-based attention model architecture.

Layer (type)	**Output Shape**	**Param #**	**Connected to**
input_layer_2 (InputLayer)	(None, 4, 1)	0	-
lstm_1 (LSTM)	(None, 4, 64)	16,896	input_layer_2[0]
dense_4 (Dense)	**(None, 4, 1)**	**65**	**lstm_1[0]**
flatten_1 (Flatten)	**(None, 4)**	**0**	**dense_4[0]**
activation (Activation)	**(None, 4)**	**0**	**flatten_1[0]**
repeat_vector (RepeatVector)	**(None, 64, 4)**	**0**	**activation[0]**
permute (Permute)	**(None, 4, 64)**	**0**	**repeat_vector[0]**
multiply (Multiply)	**(None, 4, 64)**	**0**	**lstm_1[0], permute[0]**
lambda (Lambda)	(None, 64)	0	multiply[0]
dense_5 (Dense)	(None, 64)	4,160	lambda[0]
dropout_2 (Dropout)	(None, 64)	0	dense_5[0]
dense_6 (Dense)	(None, 3)	195	dropout_2[0]

- Total params: 21,316 (83.27 KB)

- Trainable params: 21,316 (83.27 KB)
- Non-trainable params: 0 (0.00 B)

4.4 Evaluation metrics

Using metrics like recall, precision, accuracy, and F1 score, we evaluate the efficacy of every deep learning model. Model effectiveness may be gleaned from each statistic in its own unique way.

Accuracy

Accuracy evaluates the overall effectiveness of a classifier. Class designations are most advantageous when they hold equal value and are proportional.

$$\text{Accuracy} = \frac{\text{True Positives} + \text{True Negatives}}{\text{Total Samples}} \tag{4.10}$$

F1 Score

For imbalanced datasets, the F1 score is particularly helpful since it balances recall and accuracy:

$$F_1 = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \tag{4.11}$$

Recall

Recall measures a model's capacity to recognize all pertinent instances:

$$\text{Recall} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Negatives}} \tag{4.12}$$

Precision

Precision quantifies the accuracy of positive predictions, which is essential when false positives incur significant costs:

$$\text{Precision} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Positives}} \tag{4.13}$$

Confusion Matrix

This table delineates the performance of a classification model on a dataset with known true values. It facilitates the viewing of an algorithm's performance. The matrix juxtaposes the actual target values against those forecasted by the machine learning model, offering insight into the quantity of accurate and erroneous predictions categorized by each class.

4.5 Conclusion

This chapter delineated the architecture and implementation procedures of the proposed real-time stress detection system. Each stage, from data pretreatment and augmentation to the building of LSTM-based models, was meticulously crafted to improve classification performance. The technology utilizes physiological signals, specifically PPG, to accurately detect stress levels in real time. The framework guarantees resilience, scalability, and adaptability in real-world circumstances through the implementation of deep learning and ensemble approaches. The technical foundations equip the system for advanced assessment, which is further examined in the subsequent chapter.

Chapter 5

Results And Discussion

5.1 Indroduction

Tools that are used in the process of bringing a project to life are the cornerstone of every successful endeavor.This section investigates the development environment that underpins our solution and provides a road map for those individuals who want to build upon or replicate our efforts or expand upon the work that we have done that we have done.

5.2 Development environment

5.2.1 Hardware specifications

In order to implement our solution, we used:

- A laptop computer equipped with an Intel ntel(R) Core(TM) i5-5200U CPU @ 2.20GHz 2.19 GHz, 8 GB of RAM, a 256 GB SSD hard drive, and running the Windows 10 Pro operating system was utilized for pre-processing.

5.2.2 Environment and Development Tools :

Google Colaboratory :

Colaboratory, or "Colab" for short, is an interactive environment that enables users to compose and execute Python code within a web browser. It is a valuable instrument for machine-learning and deep-learning applications, as it provides free access to GPUs and TPUs and requires no configuration. In order to utilize Google Colab, you must have a Gmail account that enables you to construct and manage Colab notebooks, which are where you can write and execute your code.

Key Features and Benefits

- **Interactive Environment:** Colab provides an interactive environment similar to Jupyter Notebooks, permitting usedrs to write, execute, and debug code seamlessly.
- **Pre-installed Libraries:** TensorFlow, Keras, PyTorch, NumPy, and pandas are among the numerous essential libraries that are pre-installed. Other libraries may be installed by users through the pip-installation command.
- **Easy Data Loading:** Many essential libraries, including TensorFlow, Keras, PyTorch, NumPy, and pandas, are pre-installed. Additional libraries may be installed by users through the pip-installation command.
- **Cloud Storage Integration:** Users can store Colab Notebooks and other files by integrating with Google Drive. This facilitates effortless collaboration and sharing.

Resource Availability

- **Session Time:** Google Colab offers a continuous execution period of 12 hours per session. Data that is not saved to external storage (such as Google Drive) will be lost, and the session must be resumed after this period.
- **RAM Available Space:** The Colab provides 12 GB of RAM for each session. As a result, there is no freedom to enhance this.
- **Storage Available Space:** Google Colab offers over 30 GB of transient storage during a session, in addition to Google Drive storage. Python modules and files were stored in this area; however, they were eliminated upon the conclusion of the session.

Python Modules :

Python is a high-level programming language that is interpreted, object-oriented, and characterized by its dynamic semantics. It is an ideal choice for rapid application development and as a scripting language to integrate existing components due to its dynamic typing and binding capabilities, as well as its advanced built-in data structures. Python's syntax is straightforward and straightforward to comprehend, which contributes to its legibility and decreases the expense of program maintenance. It encourages the reuse of code and the support of modules and packages. The Python interpreter and its extensive standard library are available for free in both source and binary forms for all main platforms and can be freely disseminated.

The following modules, along with Python version 3.10.12, were employed in the Google Colab execution environment:

Libraries :

Libraries for pre-processing

- **Pandas :**Pandas is a BSD-licensed, open-source library that offers data structures and data analysis tools that are both user-friendly and high-performance for the Python programming language. This is primarily employed for the purpose of accessing and manipulating datasets in formats such as Excel and CSV files. We employ pandas to read the datasets and manipulate the data frames.
- **NumPy :**NumPy is a critical Python package that is employed for scientific computation. It offers a collection of mathematical functions to operate on these arrays, as well as support for enormous multidimensional arrays and matrices. NumPy is indispensable for linear algebra routines, random number generation, and numerical computations. NumPy was employed in our research to designate variable types, reshape arrays, and store model outputs.
- **Matplotlib :**Matplotlib is a comprehensive library that enables the development of static, animated, and interactive visualizations in Python. It offers a diverse selection of plotting functions, such as scatter plots, line plots, bar charts, histograms, and more. Matplotlib is often used in conjunction with other data analysis and machine learning libraries, such as NumPy and Pandas.

- **Scikit-learn :**It is a library for machine learning and deep learning used to partition data into training and testing sets.

Libraries for training and testing

- **Keras :**Keras is a high-level neural networks API developed in Python, designed to operate on TensorFlow, CNTK, or Theano. It is intended to facilitate rapid experimentation with deep neural networks and accommodates both convolutional and recurrent networks, along with a the amalgamation of the two. Keras is extensively utilized in deep learning and is recognized for its simplicity and user-friendliness.
- **Tensorflow :**TensorFlow is an open-source framework for machine learning created by Google. It offers an extensive ecosystem of tools, libraries, and community resources enabling developers to construct and deploy machine learning-driven applications. TensorFlow is very adept at managing extensive, intricate machine learning initiatives and is employed across a diverse array ranging from image recognition to natural language processing.
- **Sklearn.metrics :**This scikit-learn module offers a variety of metrics for the assessment of the efficacy of deep learning and machine learning models. In order to utilize functions that calculate roc-auc, recall, precision, and f1 scores, we have elected to depend on this library. Furthermore, we generate a matrix of perplexity and With this module, classification reports are generated.

Libraries for plotting

- **Seaborn :** Seaborn is a Python library for data visualization that is built upon matplotlib. It offers a sophisticated interface for creating visually appealing and instructive statistics visuals.
- **Matplotlib :**Matplotlib is a comprehensive library that enables the development of static, animated, and interactive visualizations in Python. It offers a diverse selection of plotting functions, such as scatter plots, line plots, bar charts, histograms, and more. Matplotlib is often used in conjunction with other data analysis and machine learning libraries.

5.3 Dataset for stress detection using PPG signals

A comprehensive dataset is essential for creating an efficient deep learning model for real-time stress detection.Therefore, it is crucial to identify a dataset that encompasses PPG signals from diverse subjects, including individuals of different ages and genders, who have experienced both relaxation and stressful conditions[33]. The WESAD dataset, compiled by Schmidt et al. in 2018 , contains raw photoplethysmography (PPG) signals collected from wrist-worn wearable sensors on 17 participants during scientifically-validated stressors and relaxation states. Participants specifically donned the Empatica E4 wristband and the RespiBAN chest sensor. The Empatica E4 is a wrist-mounted device equipped with photoplethysmography (PPG) sensors for assessing blood volume pulse, electrodermal activity (EDA) sensors for monitoring variations in skin conductance, accelerometers for motion detection, and infrared thermopiles for measuring skin temperature. It provides key physiological signals relevant to stress detection at a sampling rate of 64 Hz [33] suitable for collecting high- fidelity physiological signals for developing deep learning models. The RespiBAN professional system is a wireless chest-worn sensor band that contains a respiratory inductance pleth-ysmography sensor to measure breathing waveforms and detect the respiration rate [33].

Figure 5.3 Graphical depiction of the class distribution in the dataset. The pie chart (top) displays the percentage distribution of each class (Baseline, Stress, and Amusement), whilst the bar chart (bottom) represents the total number of samples for each class. This distribution reveals a class imbalance, with Stress samples markedly exceeding the others.

Signal Acquisition and Sampling Characteristics

The raw BVP signal was obtained from the Empatica E4 wearable device, which samples the signal at a frequency of 64 Hz. For each subject, the signal data was accessed from the dataset's `.pkl` files, alongside the corresponding annotations indicating the subject's experimental condition: *baseline*, *stress*, or *amusement*.

Time Alignment and Synchronization

Given that other physiological modalities in the dataset (e.g., electrodermal activity (EDA), temperature (TEMP), accelerometer (ACC), and respiration (RESP)) are sampled at different frequencies, temporal synchronization was necessary. Each BVP time series was indexed using timestamp-based alignment to enable multimodal fusion and ensure consistency across sensor streams.

Label Statistics

The dataset is annotated with three distinct emotional states:

- **Baseline (Label 0):** 195 samples
- **Stress (Label 1):** 628 samples

- **Amusement (Label 2):** 355 samples

The stress condition encompasses the highest quantity of samples, succeeded by amusement and baseline. The baseline class, possessing the least number of samples, signifies the distribution's tail. This data is a summary of the distribution of emotional states throughout the dataset.

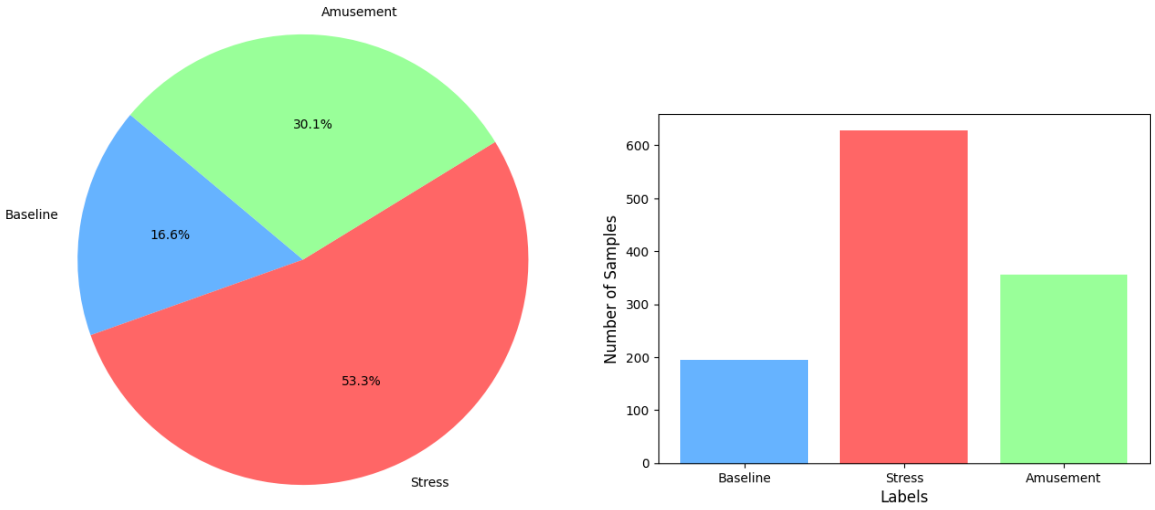


Figure 5.1: Label Distribution in WESAD Data

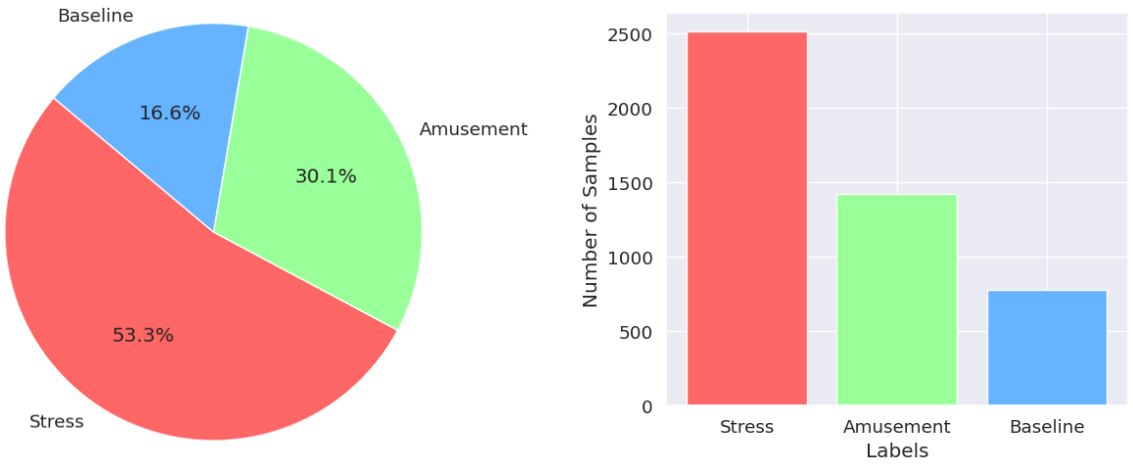


Figure 5.2: Label Distribution in WESAD Data After Augmentation

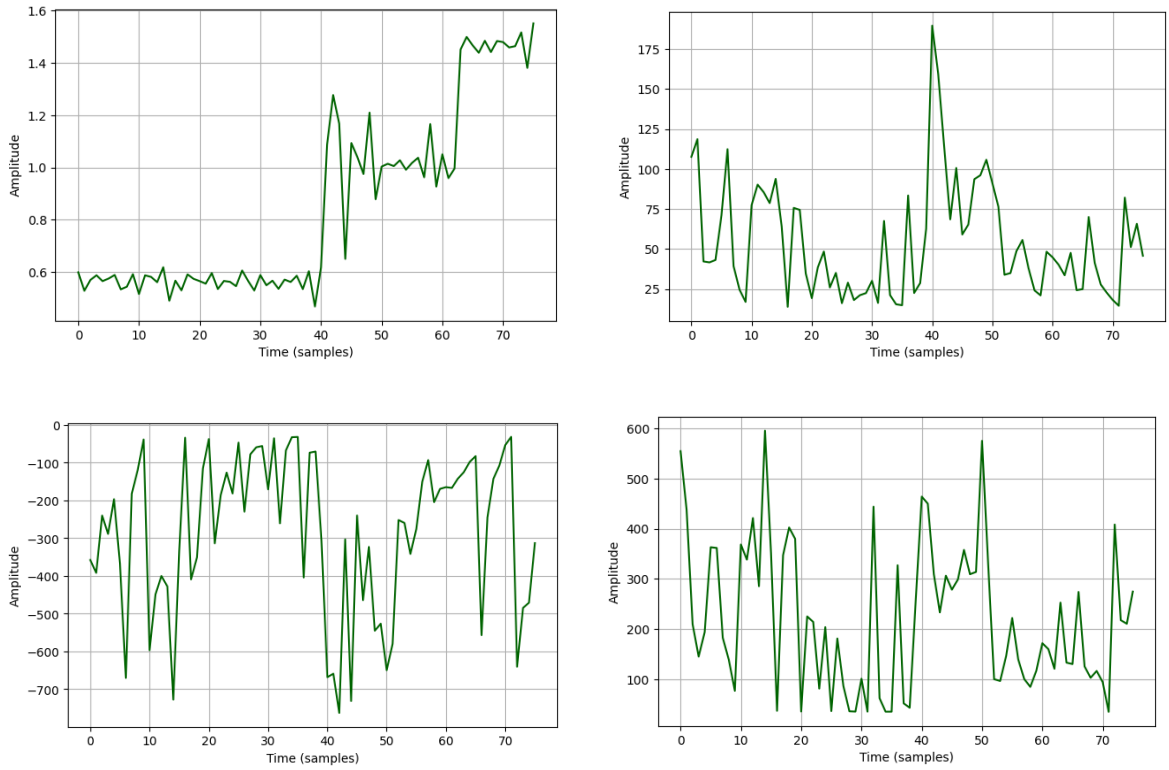


Figure 5.3: Example PPG (BVP) Signals from WESAD Subjects

5.3.1 Results :

LSTM

We conducted a comprehensive examination of the Long Short-Term Memory (LSTM) model's efficacy in categorizing emotional states from physiological signals, utilizing accuracy and loss curves alongside a confusion matrix The figures bellow shows the evaluation of the LSTM model

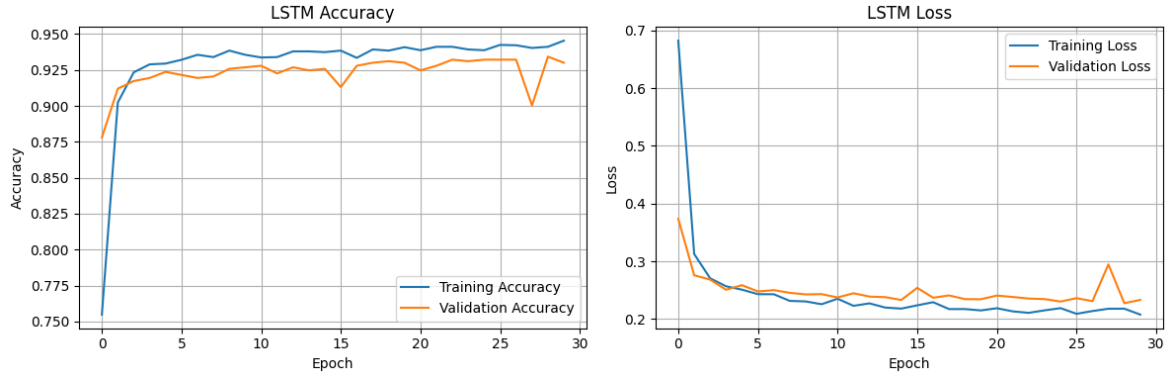


Figure 5.4: LSTM Model Training and Validation Performance Over Epochs

Performance in Training and Validation

Figure 5.4 illustrates the training and validation accuracy and loss of the LSTM model over epochs. The accuracy plot (left) clearly demonstrates that the model achieves quick convergence in the early epochs, with both training and validation accuracy stable at 90%. This signifies that the model adeptly assimilates the fundamental patterns in the training data while preserving generalization efficacy on unobserved validation data.

The loss plot (right) illustrates a steady decrease in both training and validation losses, ultimately converging, which strongly suggests effective model learning with negligible overfitting. The disparity between the two curves remains minimal, underscoring the model's capacity to generalize effectively to novel data.

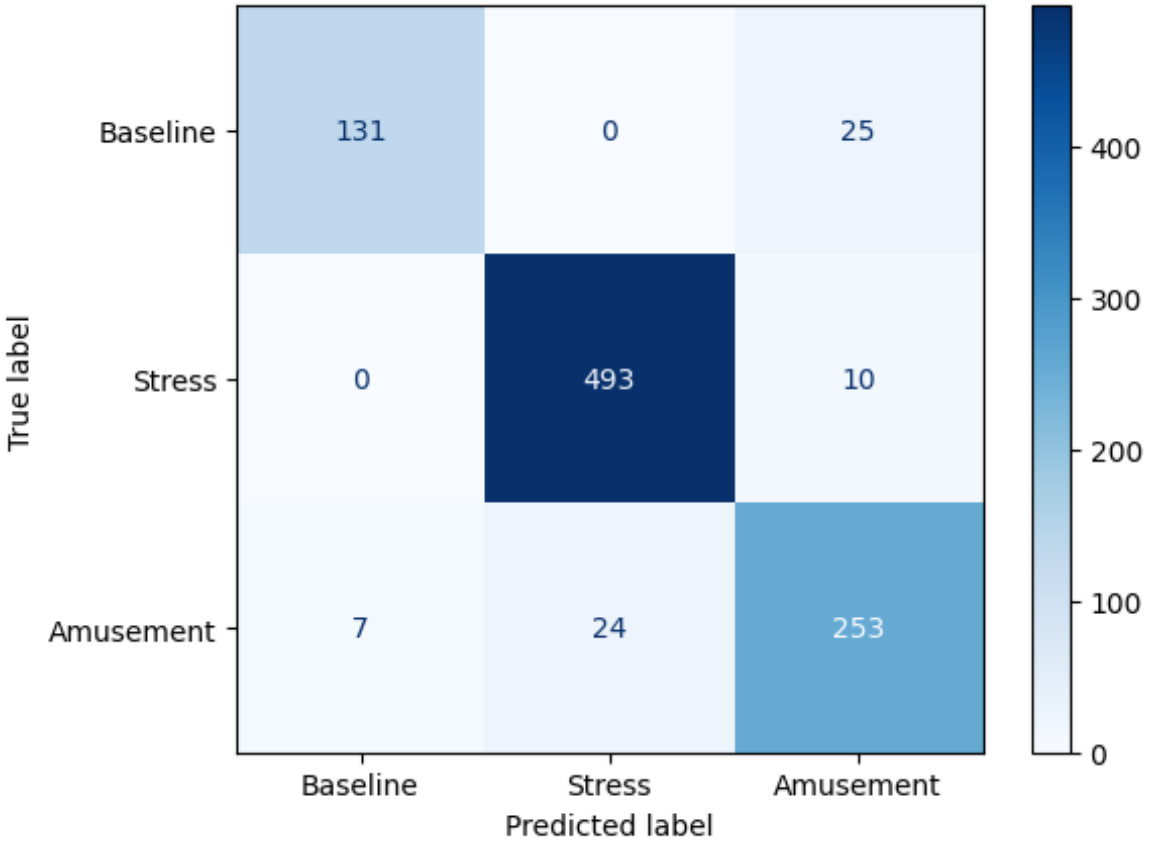


Figure 5.5: Confusion Matrix of LSTM Predictions for Baseline, Stress, and Amusement Classes

Classification Performance Using Confusion Matrix

Figure 5.5 illustrates the confusion matrix for the LSTM model's predictions across three categories: Baseline, Stress, and Amusement. The matrix discloses the subsequent insights: The baseline was accurately identified 131 times, with slight misunderstanding over the Amusement class (25 occasions).

Stress was the most reliably categorized category, with 463 true predictions with negligible misclassification.

Amusement exhibited robust classification performance, achieving 253 accurate predictions with minimal misunderstanding regarding the other two categories.

The results indicate that the LSTM model is exceptionally proficient, especially in differentiating the Stress state, which is vital for stress detection applications. Nonetheless, minor similarities between Baseline and Amusement indicate that subsequent research should investigate more sophisticated feature engineering or attention strategies to improve differentiation among emotional states.

LSTM with attention

We performed a thorough evaluation of the Long Short-Term Memory (LSTM) model enhanced with attention processes to categorize emotional states—Baseline, Stress, and Amusement—from physiological or behavioral signals. The assessment utilizes accuracy and loss trends (Figure 5.6) along with a confusion matrix (Figure 5.7) to measure performance.

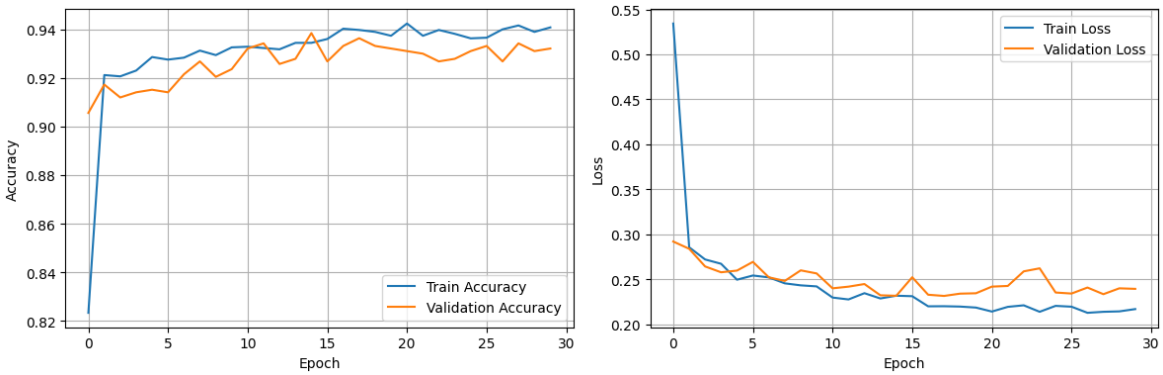


Figure 5.6: LSTM with attention Model Training and Validation Performance Over Epochs

Performance in Training and Validation

Figure 5.6 depicts the training and validation performance of the LSTM-with-attention model throughout the epochs. The accuracy curve (left) illustrates swift convergence during the initial epochs, with both training and validation measures stabilizing at a high performance level, signifying effective learning of temporal patterns in the data. The strong correspondence between the two curves further validates the model's robustness, as it preserves generalization ability without succumbing to overfitting on the training data.

The loss curve (right) demonstrates a consistent decrease for both training and validation, ultimately stabilizing at a minimal amount. The steady decrease, along with the minimal disparity between the curves, highlights the model's optimization stability and its capacity to generalize to unfamiliar data. The attention mechanism probably

enhances this efficiency by dynamically assigning weights to pertinent temporal features during training.

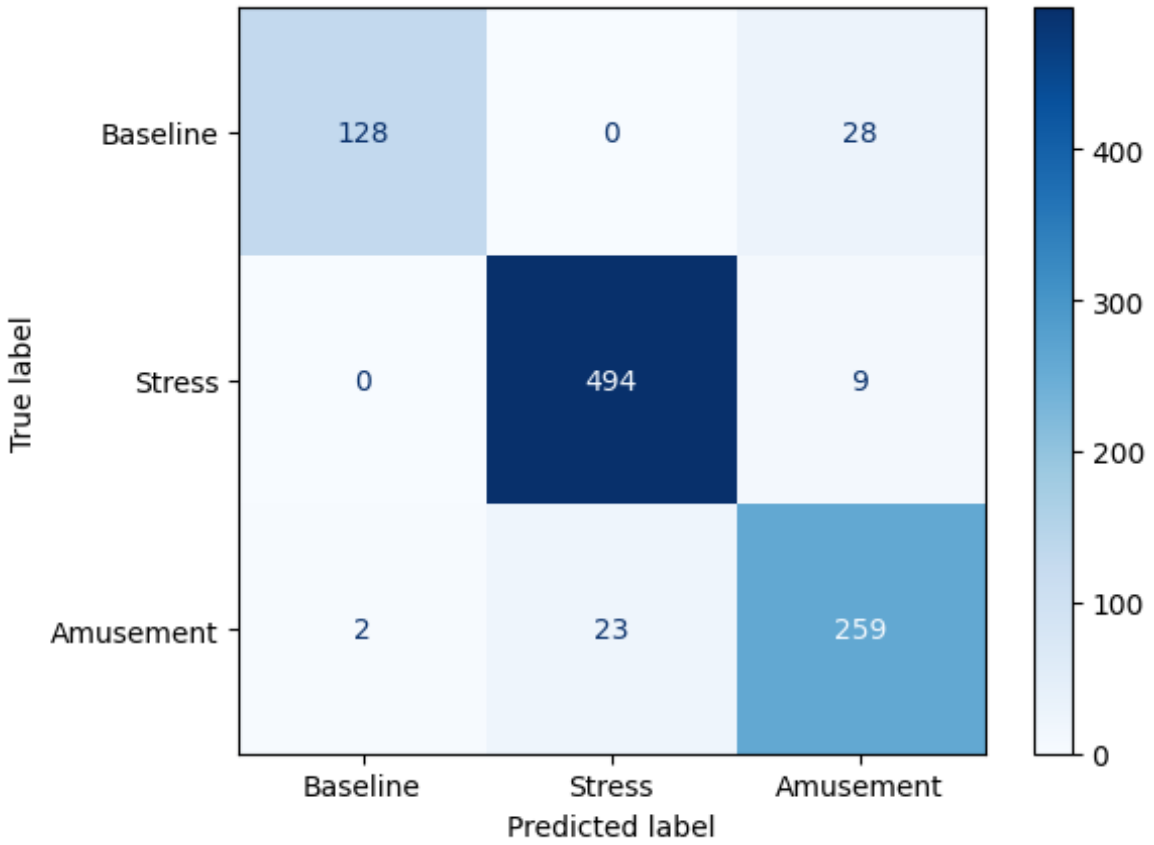


Figure 5.7: Confusion Matrix - LSTM with Attention

Classification Performance Using Confusion Matrix

Figure 5.7 depicts the confusion matrix for the predictions of the LSTM-with-attention model across three emotional states: Baseline, Stress, and Amusement. The matrix presents the subsequent insights:

Baseline: Attained 128 accurate classifications (true positives), although demonstrated confusion with Amusement (2 misclassifications) and Stress (0 misclassifications). This indicates that neutral moods may possess nuanced physiological or temporal characteristics akin to Amusement, hence confounding their differentiation.

Stress: Exhibited robust classification effectiveness with 494 true positives and merely 9 misclassifications, all categorized as Amusement. The model's accuracy in identifying stress underscores its appropriateness for applications such as stress monitoring, where elevated recall is essential.

Amusement demonstrated strong performance with 293 correct predictions, while there was slight misunderstanding with Baseline, resulting in 23 misclassifications. The attention mechanism probably improved the differentiation of Amusement's unique high-arousal patterns.

Training Configuration

The model was compiled using the Adam optimizer and categorical cross-entropy loss function, which is suitable for multi-class classification. Accuracy was used as the performance metric.

- **Optimizer**: Adam
- **Loss Function**: Categorical Crossentropy
- **Metric**: Accuracy
- **Epochs**: 30
- **Batch Size**: 32
- **Validation Set**: 20% of the dataset

The input data was reshaped to a format of `(samples, 4, 1)`, and the labels were one-hot encoded to `(samples, 3)`.

Table 5.1: Comprehensive Comparison of Machine Learning and Deep Learning Classifiers

Model	**Method**	**Accuracy**	**Macro F1**	**Time (s)**
RandomForest	RandomizedSearchCV	0.9396	0.93	23.07
KNN	RandomizedSearchCV	0.8134	0.78	1.05
Logistic Regression	RandomizedSearchCV	0.9332	0.92	3.34
DecisionTree	RandomizedSearchCV	0.9300	0.91	0.92
CNN	—	0.9200	0.91	0.27
LSTM	—	**0.9300**	**0.91**	**0.06**
LSTM with Attention	—	0.9323	0.92	0.09
Ensemble (CNN + LSTM)	Averaged predictions	0.9300	0.92	0.1

Discussion

Table 5.1. The comparative performance table offers a thorough assessment of different machine learning (ML) and deep learning (DL) models for stress detection utilizing characteristics derived from the WESAD dataset. The models are evaluated according to four primary criteria: hyperparameter tuning methodology, classification accuracy, macro F1-score, and prediction time per sample.

The Random Forest model achieved the best accuracy (93.96%) and largest F1-score (0.93), indicating strong classification performance. Nonetheless, this resulted in a **considerably prolonged inference time (23.07 seconds per sample)**, rendering it less appropriate for real-time stress detection applications. The dependence on ensemble decision trees probably resulted in both elevated accuracy and increased computing demands.

Conversely, **K-Nearest Neighbors (KNN)** attained the **lowest accuracy (81.34%)** and macro F1-score (0.78), accompanied by a moderate prediction time (1.05 s/sample). This underperformance may arise from KNN's susceptibility to noisy data and its comparatively basic framework in high-dimensional physiological datasets.

Logistic Regression achieved an accuracy of 93.32% and a macro F1-score of 0.92, with a prediction time of 3.34 seconds. This suggests that linear models can effectively identify important discriminative patterns in the data, particularly when combined with suitable regularization and feature selection.

Decision Tree models demonstrated a commendable equilibrium, attaining 93.00% accuracy and a 0.91 macro F1 score, while exhibiting superior efficiency compared to Logistic Regression (0.92 s/sample). Nevertheless, individual decision trees are susceptible to overfitting, potentially constraining their generalizability despite yielding robust results.

Concerning deep learning models, both CNN and LSTM attained robust outcomes. LSTM marginally surpassed CNN in accuracy (93.00% vs. 92.00%) while exhibiting the lowest prediction time (0.06 s/sample) among all models, rendering it exceptionally advantageous for real-time applications. This indicates that LSTM's capacity to capture temporal correlations in physiological signals such as BVP is especially beneficial for stress identification.

The ensemble of CNN and LSTM, through averaged predictions, preserved the same accuracy as LSTM (93.00%) while enhancing the macro F1-score to 0.92, signifying improved class-wise equilibrium. The time (**0.1 s/sample**) is within acceptable real-time parameters, affirming its appropriateness for implementation in mobile or wearable contexts.

The results indicate that LSTM and the CNN+LSTM ensemble provide the optimal balance between performance and real-time applicability. Although conventional machine learning models such as Random Forest and Logistic Regression attain elevated accuracies, their inference durations constrain their usefulness. The findings highlight the superiority of deep learning, especially recurrent architectures such as LSTM, for real-time physiological stress detection systems.

5.4 Comparison of Proposed Model with Related Works

Table 5.2: Comparison of Proposed Model with Related Works in rPPG-based Stress Detection

Study / Method	**Approach / Methodology**	**Dataset Used**	**Accuracy (%)**	**Key Features / Notes**
Sabour et al. [30]	rPPG-based stress estimation system	Not specified	85.48	Non-contact stress estimation using HRV features from video
Benezeth et al. [6]	rPPG-based HRV estimation for emotional states detection	Not specified	Correlation only	Showed strong correlation between HRV and emotions (no specific accuracy reported)
Hybrid DL Network (UBFC-Phys)	Deep Learning (LSTM) for rPPG signal and HR estimation	UBFC-Phys	95.83	Uses hybrid DL model for stress detection, high precision and efficacy
Proposed Model (This Work)	**LSTM on BVP features extracted from WESAD**	**WESAD**	**93.00**	**Real-time stress detection using PPG(BVP) features**
Traditional ML methods (e.g., MVICAR) [10]	Statistical & ML algorithms for HRV analysis	Wearable devices	∼80-90	Applied on HRV data, nonlinear modeling (depends on method)
EDA-based methods [12] [32]	EDA signal with biofeedback via mobile app or wearables	Wearable devices	Variable	Continuous EDA monitoring; accuracy influenced by ambient conditions

Discussion:
Table 5.2 summarizes a comparison between the proposed LSTM-based model and several state-of-the-art rPPG or physiological signal-based stress detection systems. Sabour et al.[30] and Benezeth et al.[6] explored non-contact stress estimation methods using rPPG signals derived from facial video input, achieving moderate accuracy or

strong HRV-emotion correlations, respectively. However, their approaches often lacked real-time applicability and used unspecified datasets, limiting reproducibility.

Hybrid deep learning frameworks, such as the one used in UBFC-Phys, leveraged advanced LSTM models to achieve high accuracy (95.83%), showcasing the benefit of using deep temporal modeling. Nevertheless, their reliance on non-wearable systems and a different dataset (UBFC-Phys) makes direct comparison difficult.

In contrast, the proposed model employs a contact-based method using the publicly available WESAD dataset, achieving a competitive accuracy of 93.00%. It uses LSTM networks trained on BVP features, offering a practical solution for real-time stress detection in wearable scenarios.

Other approaches, such as MVICAR[10], rely on classical machine learning models applied to HRV signals with variable accuracy depending on the method and dataset. EDA-based systems [12, 32] are also common in mobile applications, but their accuracy can be affected by environmental factors and sensor placement.

These findings emphasize the relevance of the proposed approach as a strong alternative, balancing real-time performance, reliability, and sensor feasibility for wearable stress detection applications.

5.5 Conclusion

This chapter presents an analysis and interpretation of the data derived from the employed machine learning and deep learning models for stress detection. The LSTM with an attention mechanism exhibited enhanced performance relative to baseline models, validating its capacity to efficiently capture temporal correlations in physiological inputs. The application of data augmentation and hyperparameter optimization significantly improved model accuracy and generalization. Furthermore, the comparison with analogous studies underscored the competitiveness and dependability of our suggested methodology in real-time stress classification. These favorable results confirm the efficacy of the proposed method and establish a foundation for practical implementation and more research.

General Conclusion

This work introduces a thorough methodology for creating a real-time stress detection system utilizing the WESAD dataset. By concentrating on physiological indicators like Photoplethysmography PPG (BVP) and employing several machine learning and deep learning methodologies—including CNN, LSTM, and ensemble models—we established the viability and efficacy of automated stress recognition. The findings underscored the efficacy of integrating sophisticated signal processing with resilient classification techniques to attain elevated accuracy and real-time functionality.

Additionally, our assessment encompassed both model efficacy and computational efficiency, which are critical for real-time applications. Evaluating various strategies using metrics like accuracy, prediction time per sample, and model size yields important insights into the trade-offs between complexity and performance. The results confirm the feasibility of implementing these technologies in real-world applications for mental health surveillance and occupational wellness.

Future Work

To augment the system's capabilities, many avenues for future research are proposed:

- **Multimodal Signal Integration:** Integrating BVP with additional physiological signals such as EDA, EMG, or breathing to enhance classification robustness.
- **Real-World Testing:** Implementing the system on mobile or wearable platforms for real-time surveillance in unregulated areas.
- **Personalized Models:** Creating subject-specific models or adaptive learning systems to address individual variations in stress response.
- **Explainability and Interpretability:** Incorporating explainable AI (XAI) techniques to enhance transparency and trust in model predictions.
- **Dataset Expansion:** Implementing the methodology on larger or more heterogeneous datasets to enhance generalizability.
- **Lightweight Model Optimization:** Utilizing quantization, pruning, or neural architecture search (NAS) for the deployment of deep learning models on resource-limited devices.

These improvements seek to connect laboratory research with practical applications, guaranteeing the efficacy and dependability of stress detection systems in daily use.

Bibliography

[1] L. Abbas, S. Samy, R. Ghazal, A. M. Eldeib, and S. H. ElGohary. Contactless vital signs monitoring for public health welfare. In *Proceedings of the 2021 9th International Japan-Africa Conference on Electronics, Communications, and Computations (JAC-ECC)*, pages 183–186. IEEE, 2021.

[2] Robert B. Aird and Y. Gastaut. Occipital and posterior electroencephalographic rhythms. *Electroencephalography and Clinical Neurophysiology*, 11(4):637–656, November 1959.

[3] Anishnama. Understanding lstm: Architecture, pros and cons, and implementation, 2023. Accessed: 2023-04-28.

[4] Aamir Arsalan, Syed Muhammad Anwar, and Muhammad Majid. Mental stress detection using data from wearable and non-wearable sensors: A review. *CoRR*, abs/2202.03033, 2022.

[5] Almudena Bartolomé-Tomás, Roberto Sánchez-Reolid, Alicia Fernández-Sotos, José Miguel Latorre, and Antonio Fernández-Caballero. Arousal detection in elderly people from electrodermal activity using musical stimuli. *Sensors*, 20(17):4788, 2020.

[6] Y. Benezeth, S. Bobbia, K. Nakamura, R. Gomez, and J. Dubois. Probabilistic signal quality metric for reduced complexity unsupervised remote photoplethysmography. In *Proceedings of the 2019 13th International Symposium on Medical Information and Communication Technology (ISMICT)*, pages 1–5. IEEE, 2019.

[7] biosignalsplux. Eda biosignalsplux. `https://support.pluxbiosignals.com/wp-content/uploads/2021/11/Electrodermal_Activity_EDA_Datasheet.pdf`, 2022. [Online; accessed 26-April-2025].

[8] biosignalsplux. Eeg biosignalsplux. `https://support.pluxbiosignals.com/wp-content/uploads/2021/11/Electroencephalography_EEG_Datasheet.pdf`, 2022. [Online; accessed 26-April-2025].

[9] biosignalsplux. Resp biosignalsplux. `https://support.pluxbiosignals.com/wp-content/uploads/2021/11/Respiration_PZT_Datasheet.pdf`, 2022. [Online; accessed 26-April-2025].

[10] P. Caminal, F. Sola, P. Gomis, E. Guasch, A. Perera, N. Soriano, and L. Mont. Validity of the polar v800 monitor for measuring heart rate variability in mountain

running route conditions. *Eur. J. Appl. Physiol.*, 118:669–677, 2018. [CrossRef] [PubMed].

[11] Denisse Castaneda, Aibhlin Esparza, Mohammad Ghamari, Cinna Soltanpur, and Homer Nazeran. A review on wearable photoplethysmography sensors and their potential future applications in health care. *International Journal of Biosensors Bioelectronics*, 4(4):195–202, 2018.

[12] Y. Chen, M. Rao, K. Feng, and G. Niu. Modified varying index coefficient autoregression model for representation of the nonstationary vibration from a planetary gearbox. *IEEE Trans. Instrum. Meas.*, 72:1–12, 2023. [CrossRef].

[13] Y. C. Cheng, T. I. Chou, F. I. Indikawati, S. Winiarti, A. Dahlan, R. Selatan, and D. Yogyakarta. Stress detection from multimodal wearable sensor data. In *IOP Conference Series: Materials Science and Engineering*, volume 771, page 012028. IOP Publishing, 2020.

[14] N.S. Chudy. Testing of wrist-worn-fitness-tracking devices during cognitive stress: A validation study, 2017.

[15] Rohan Gandhi. Support vector machine — introduction to machine learning algorithms, June 2018. Retrieved from Medium.

[16] GeeksforGeeks. Semi-supervised learning in ml, May 2024. Accessed: 2025-04-05.

[17] Giorgos Giannakakis, Dimitris Grigoriadis, Katerina Giannakaki, Olympia Simantiraki, Alexandros Roniotis, and Manolis Tsiknakis. Review on psychological stress detection using biosignals. *IEEE Transactions on Affective Computing*, 13(1):440–460, 2022.

[18] D. Giles, N. Draper, and W. Neil. Validity of the polar v800 heart rate monitor to measure rr intervals at rest. *Eur. J. Appl. Physiol.*, 116:563–571, 2016. [CrossRef].

[19] Alexander S. Gillis, Cameron Hashemi-Pour, and Joseph M. Carew. What is reinforcement learning?, 2024.

[20] Noor Halim, Khairul Sidek, and H. Mansor. Stress recognition using photoplethysmogram signal. *Indonesian Journal of Electrical Engineering and Computer Science*, 8:495–501, 2017.

[21] M. Hassan, A. Malik, D. Fofi, N. Saad, and F. Meriaudeau. Novel health monitoring method using an rgb camera. *Biomedical Optics Express*, 8(11):4838–4854, 2017. [PubMed].

[22] M. Hassan, A. Malik, D. Fofi, N. Saad, and F. Meriaudeau. Novel health monitoring method using an rgb camera. *Biomed. Opt. Express*, 8:4838–4854, 2017. [CrossRef] [PubMed].

[23] Keiko Ikeda, Kiyoshi Kawakami, Hiroshi Onimaru, Yasumasa Okada, Shigefumi Yokota, Naohiro Koshiya, Yoshitaka Oku, Makito Iizuka, and Hidehiko Koizumi. The respiratory control mechanisms in the brainstem and spinal cord: integrative

views of the neuroanatomy and neurophysiology. *The Journal of Physiological Sciences: JPS*, 67(1):45–62, January 2017.

[24] R. J. Lee, S. Sivakumar, and K. H. Lim. Review on remote heart rate measurements using photoplethysmography. *Multimedia Tools and Applications*, pages 1–30, 2023.

[25] B. Lokendra and G. Puneet. And-rppg: A novel denoising-rppg network for improving remote heart rate estimation. *Computers in Biology and Medicine*, 141:105146, 2022. [PubMed].

[26] S. Mahato and S. Paul. Detection of major depressive disorder using linear and non-linear features from eeg signals. *Microsystem Technologies*, 25(3):1065–1076, March 2019.

[27] S. Mahato and S. Paul. Classification of depression patients and normal subjects based on electroencephalogram (eeg) signal using alpha power and theta asymmetry. *Journal of Medical Systems*, 44(1):28, January 2020.

[28] D. McDuff. Camera measurement of physiological vital signs. *ACM Computing Surveys*, 55:1–40, 2023.

[29] J. Murel. Ensemble learning: A machine learning approach. *Journal of AI Research*, 32(4):123–135, 2024.

[30] R. M. Sabour, Y. Benezeth, P. De Oliveira, J. Chappe, and F. Yang. Ubfc-phys: A multimodal database for psychophysiological studies of social stress. *IEEE Transactions on Affective Computing*, 14:622–636, 2021.

[31] R.M. Sabour, Y. Benezeth, P. De Oliveira, J. Chappe, and F. Yang. UBFC-Phys: A Multimodal Database for Psychophysiological Studies of Social Stress. *IEEE Trans. Affect. Comput.*, 14:622–636, 2021. [CrossRef].

[32] L. Salahuddin, J. Cho, M.G. Jeong, and D. Kim. Ultra short term analysis of heart rate variability for monitoring mental stress in mobile settings. In *Proceedings of the 2007 29th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pages 4656–4659, Lyon, France, August 2007. 22–26 August 2007.

[33] Phillip Schmidt, Attila Reiss, Robert Duerichen, Christian Marberger, and Kristof Van Laerhoven. Introducing WESAD, a multimodal dataset for wearable stress and affect detection. In *Proceedings of the 20th ACM International Conference on Multimodal Interaction (ICMI)*, pages 400–408, New York, NY, USA, 2018. ACM.

[34] V. Selvaraju, N. Spicher, J. Wang, N. Ganapathy, J. M. Warnecke, S. Leonhardt, R. Swaminathan, and T. M. Deserno. Continuous monitoring of vital signs using cameras: A systematic review. *Sensors*, 22(11):4097, 2022.

[35] M.M. Shahid, G.E. Agada, M. Kayyali, I.K. Ihianle, and P. Machado. Towards enhanced well-being: Monitoring stress and health with smart sensor systems.

In *Proceedings of the 2023 International Conference Automatics and Informatics (ICAI)*, pages 432–437, Varna, Bulgaria, October 2023. 5–7 October 2023.

[36] S. A. Singh, P. K. Gupta, M. Rajeshwari, and T. Janumala. Detection of stress using biosensors. *Materials Today: Proceedings*, 5(10):21003–21010, 2018.

[37] SitePoint Team. Random forest algorithm in machine learning, January 2025. Accessed: 2025-01-27.

[38] R. Sánchez-Reolid, A. Martínez-Rodrigo, M.T. López, and A. Fernández-Caballero. Deep support vector machines for the identification of stress condition from electrodermal activity. *Int. J. Neural Syst.*, 30(2050031), 2020. [CrossRef] [PubMed].

[39] G. Tanev, D.B. Saadi, K. Hoppe, and H.B. Sorensen. Classification of acute stress using linear and non-linear heart rate variability analysis derived from sternal ecg. In *Proceedings of the 2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pages 3386–3389, Chicago, IL, USA, August 2014. 26–30 August 2014.

[40] Larisa G. Tereshchenko and Mark E. Josephson. Frequency content and characteristics of ventricular conduction. *Journal of Electrocardiology*, 48(6):933–937, 2015.

[41] Wikipedia contributors. Elettrocardiogramma — wikipedia, l'enciclopedia libera, 2023. [Online; accessed 14-March-2023].

[42] Z. Yang, H. Wang, and F. Lu. Assessment of deep learning-based heart rate estimation using remote photoplethysmography under different illuminations. *IEEE Transactions on Human-Machine Systems*, 52:1236–1246, 2022.

[43] Z. Yu, X. Li, and G. Zhao. Facial-video-based physiological signal measurement: Recent advances and affective applications. *IEEE Signal Processing Magazine*, 38:50–58, 2021.

Appendix

CNN

The figures bellow shows the evaluation of the CNN model

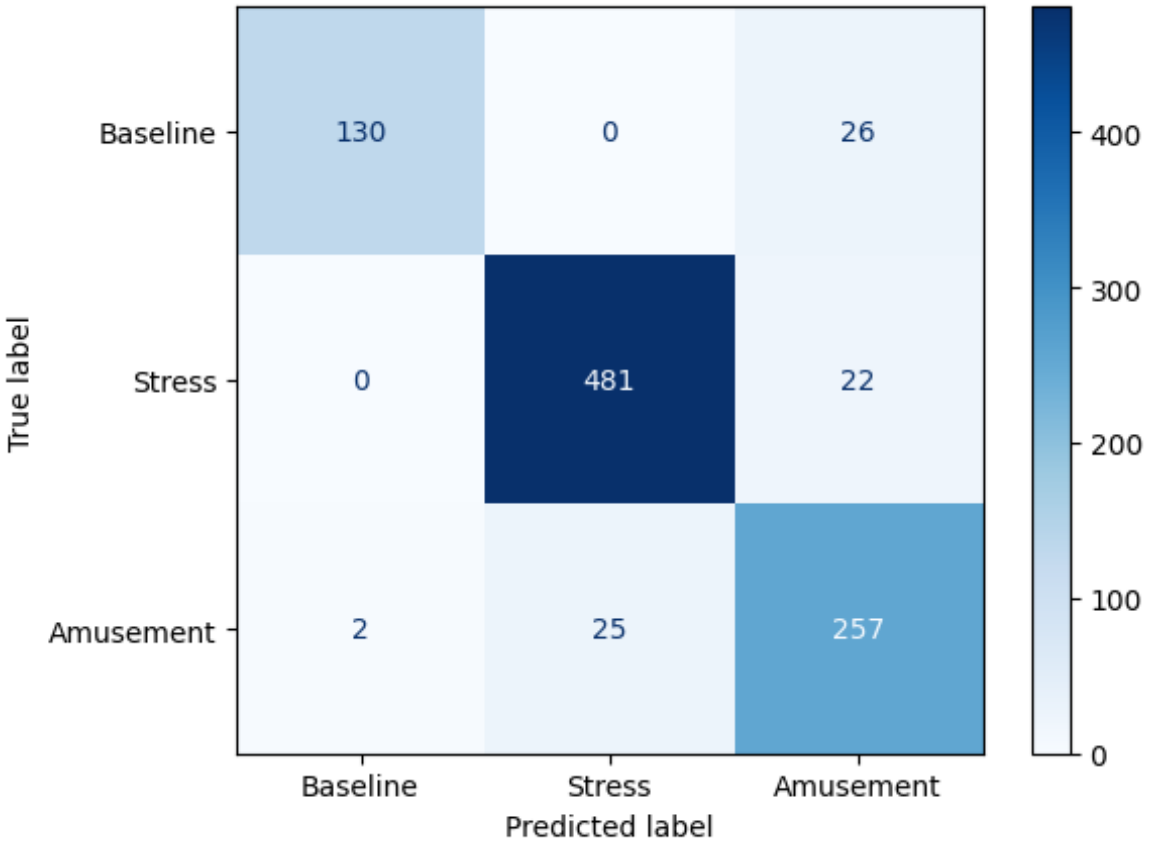


Figure 5.8: Confusion Matrix of CNN Predictions for Baseline, Stress, and Amusement Classes

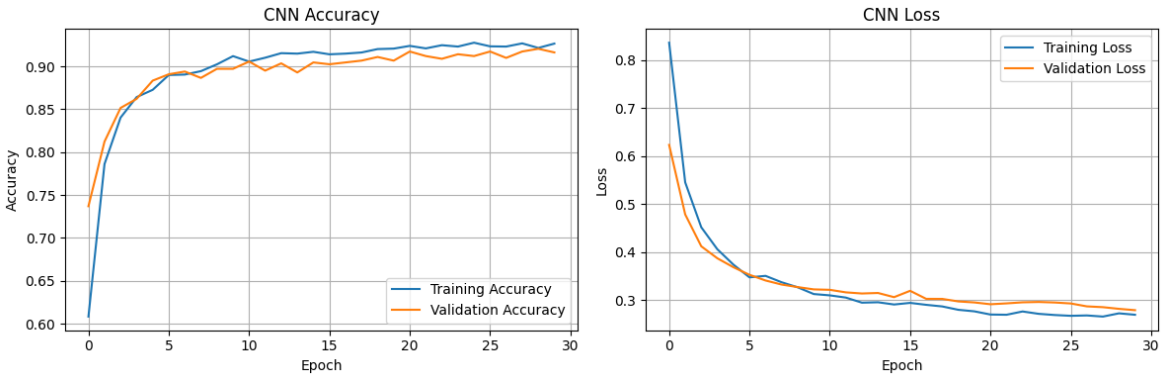


Figure 5.9: CNN Model Training and Validation Performance Over Epochs

CNN+LSTM

The figures bellow shows the evaluation of the CNN + LSTM model

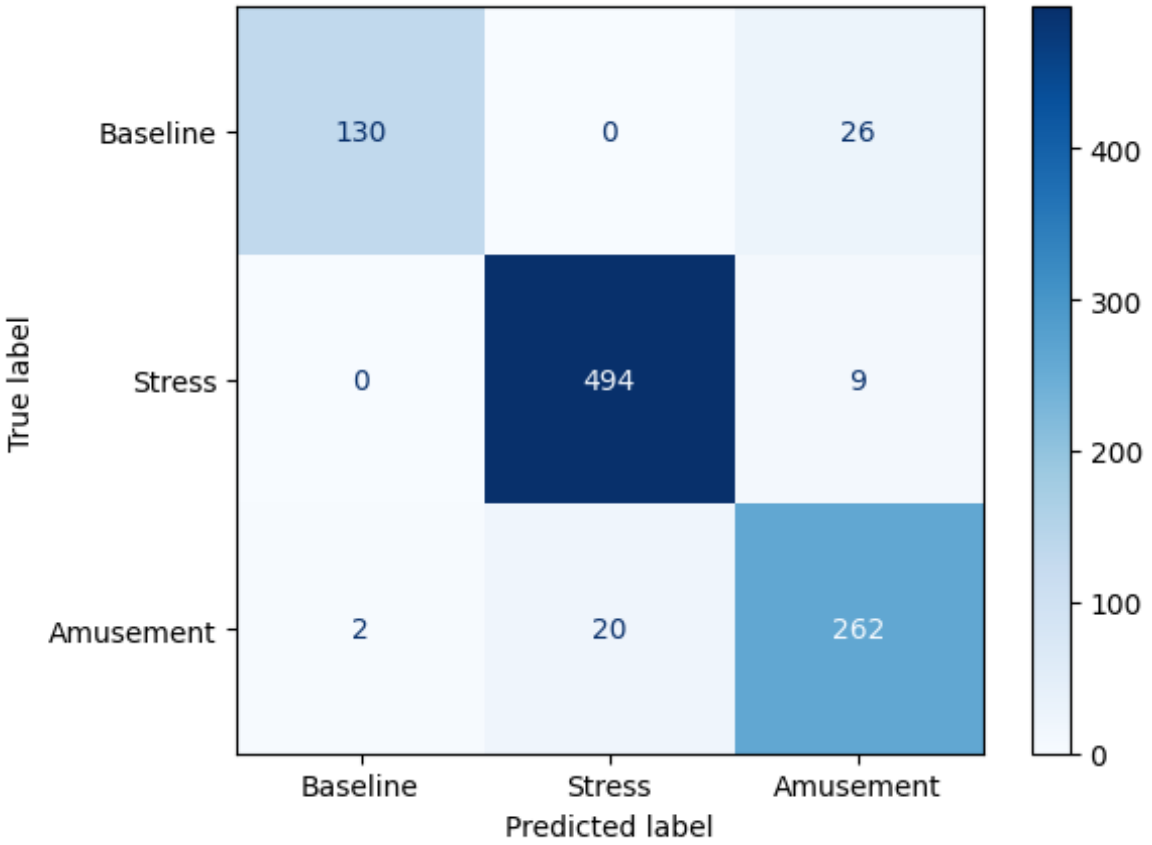


Figure 5.10: Confusion Matrix - Ensemble (CNN + LSTM)

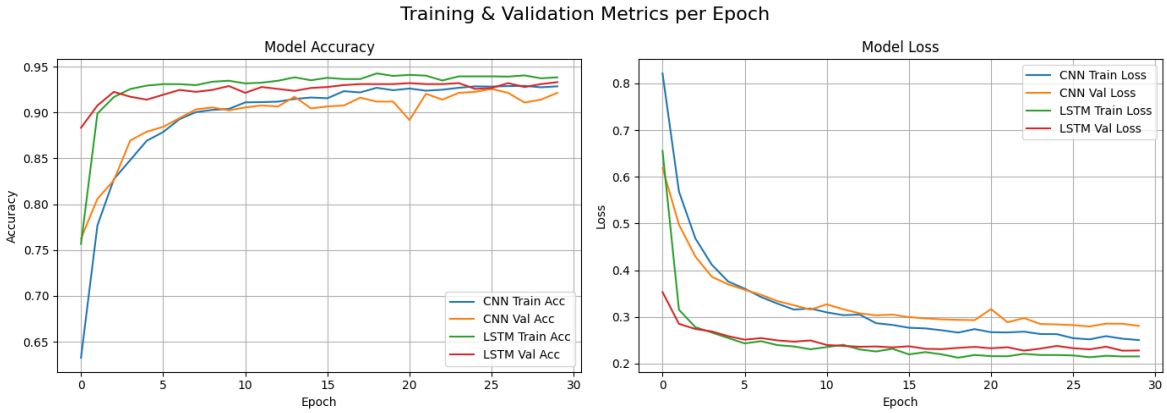


Figure 5.11: Training Validation Metrics per Epoch