PREDICTION OF STRESS LEVEL BASED ON PHYSICAL ACTIVITY USING MACHINE LEARNING

BY

ISRAK JAHAN LUBA
ID: 201-15-13855

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

Supervised By

Faria Nishat Khan

Lecturer

Department of CSE

Daffodil International University

Co-Supervised By

Mr. Md. Sazzadur Ahamed

Assistant Professor

Department of CSE

Daffodil International University

DAFFODIL INTERNATIONAL UNIVERSITY

DHAKA, BANGLADESH

JULY 2024

APPROVAL

This Project titled **"PREDICTION OF STRESS LEVELS BASED ON PHYSICAL ACTIVITY USING MACHINE LEARNING"**, submitted by ISRAK JAHAN LUBA to the Department of Computer Science and Engineering, Daffodil International University, has been accepted as satisfactory for the partial fulfillment of the requirements for the degree of B.Sc. in Computer Science and Engineering and approved as to its style and contents. The presentation has been held on **13 July, 2024.**

BOARD OF EXAMINERS

(Narayan Ranjan Chakraborty) **Associate Professor & Associate Head** Department of CSE Faculty of Science & Information Technology Daffodil International University	**Board Chairman**
(Md. Sadekur Rahman) **Assistant Professor** Department of CSE Faculty of Science & Information Technology Daffodil International University	**Internal Examiner 1**
(Mr. Tanvirul Islam) **Lecturer** Department of CSE Faculty of Science & Information Technology Daffodil International University	**Internal Examiner 2**
(Dr. Md. Manowarul Islam) **Associate Professor** Department of CSE Jagannath University	**External Examiner**

DECLARATION

I hereby declare that, this project has been done by me under the supervision of **Faria Nishat Khan, Lecturer, Department of CSE,** Daffodil International University. I also declare that neither this project nor any part of this project has been submitted elsewhere for the award of any degree or diploma.

Supervised by:

13.7.24

Faria Nishat Khan
Lecturer
Department of CSE
Daffodil International University

Co-Supervised by:

Mr. Md. Sazzadur Ahamed
Assistant Professor
Department of CSE
Daffodil International University

Submitted by:

Israk Jahan

Israk Jahan Luba
ID: 201-15-13855
Department of CSE
Daffodil International University

ACKNOWLEDGEMENT

First, I express my heartiest thanks and gratefulness to almighty God for His divine blessing makes it possible to complete the final year project/internship successfully.

I am really grateful and wish my profound indebtedness to the **supervision of Faria Nishat Khan, Lecturer,** Department of CSE Daffodil International University, Dhaka. Deep Knowledge & keen interest of my supervisor in the field of "*Machine Learning*" to carry out this project. Her endless patience, scholarly guidance, continual encouragement, constant and energetic supervision, constructive criticism, valuable advice, reading many inferior drafts and correcting them at all stages have made it possible to complete this project.

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

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

Finally, I must acknowledge with due respect the constant support and patients of my parents.

ABSTRACT

Stress has become a major risk factor in today's world, which affects different people from different demographics. This research uses a large dataset with various physiological and behavioral factors to present a novel method for detecting stress. This study will analyze 11 different variables, such as physical activity level, gender, age, sleep quality, and cardiovascular health markers etc. to build a robust model that can reliably detect and predict stress levels. My goal in this study is to create reliable models for stress detection through the analysis of several features and the use of cutting-edge machine learning methods. Stress is a major public health issue that has a big impact on people's wellbeing. Because subjective self-reporting is frequently used in traditional stress assessment methods, precise and objective measurement methods are crucial. In order to precisely identify, anticipate, and forecast stress levels based on a variety of characteristics, such as physiological, behavioral, and contextual aspects, I suggest a data-driven method. I investigate the efficiency of various machine learning models in identifying intricate patterns and relationships in the data, such as Decision Tree, Random Forests, and Gradient Boosting Machines (GBM). To extract pertinent data and improve model performance, feature engineering and selection strategies are used. I determine the most useful features for stress detection and evaluate each model's predictive ability through rigorous testing and cross-validation. In addition, I examine the underlying trends and connections between the stressors and traits that have been revealed, offering important new understandings into the variables affecting stress levels. These realizations can direct the creation of customized treatments that are suited to each person's particular requirements, enabling efficient stress management techniques. This study adds to the body of knowledge in the field of stress detection and mitigation by bridging the theoretical and practical divide. This study adds to the body of knowledge in the field of stress detection and mitigation by bridging the theoretical and practical divide. In this work, my accuracy rate is 95%.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

Introduction

1.1 Introduction

Stress has become a major risk factor in today's world. Half of the population is now affected by stress. Due to this stress, people are getting sick and the risk of life is increasing. Half of the young generation is now suffering from depression and other illnesses due to stress. This study examines comprehensive stress detection strategies using a robust dataset covering some of independent factors including heart rate, blood pressure, gender, age, sleep quality, BMI and physical activity level. Recognizing the complex nature of stress, research aims to dissect and identify the subtle connections between these different factors, examining how they collectively contribute to a person's experience of stress. individual. Through, systematic analysis using selective machine learning models, my goal is to identify characteristics that indicate stress and find correlations, thereby improving our understanding of its underlying cause. In today's fast-paced society, stress has become a prevalent occurrence that affects people from various demographic backgrounds. This study presents a novel stress detection technique using a big dataset with several physiological and behavioral parameters. Eleven variables, including age, gender, physical activity level, sleep quality, and markers of cardiovascular health, will be analyzed in this study in order to create a powerful model that can precisely detect and forecast stress levels. By carefully analyzing the information, the research aims to offer evidence-based suggestions that transcend general strategies and take into account the individual differences in stress experiences. Additionally, this study recognizes that stress is a dynamic phenomenon in the modern world that is impacted by changes in work conditions, lifestyle habits, and technology improvements. The study intends to provide insights that are not only scientifically sound but also pertinent and adjustable to the difficulties of the modern period by tackling these modern elements.

My vision is not only to enhance general health and reduce stress but also to improve stress management and promote general health, which further enhances the benefits. I have tried to work on activities that show stress levels. These activities have created a dataset with some raw data and some fake data. And tested several models using machine learning to get the best assurance.

1.2 Motivation

Motivation

My motivation behind this research is to address the growing global burden of stress-related illnesses and the urgent need for treatment. I investigate a wide range of datasets that may influence stress levels in an attempt to uncover hidden information that may guide more effective preventive and treatment strategies. By harnessing the power of machine learning to accomplish the above, I aim to create cutting-edge solutions that enable people to proactively monitor and manage their stress levels. Ultimately, I hope that my work will help bridge the gap between modern technological innovation and conventional medical practice, contributing to the rapidly developing field of digital health.

- Stress levels can be detected by other physical activities in addition to measuring devices. So, I worked on those activities to focus on those activities to reduce stress and be alert.
- I emphasized how different physical activities are for patients and common people and how important it is to know each activity.
- I can easily take a dataset and build some models with machine learning to get fast and accurate and robust results.

Objectives

- To Develop robust models for stress detection through analysis of various features.
- To Accurately identify, predict, and forecast stress levels using advanced machine learning algorithms.
- To Analyze underlying patterns and correlations between factors and stressors.
- To Provide valuable insights to guide stress reduction strategies and treatments.

1.3 Rationale of the study

Most stress assessment techniques used today rely on self-reporting, which is biased and vulnerable to subjectivity. In order to provide more accurate and trustworthy assessments of stress levels and to enable prompt responses and support, objective metrics are required. People of all demographic backgrounds are affected by stress, which is becoming more and more acknowledged as a serious public health issue. The damaging effects of stress on one's physical and mental health highlight how crucial it is to provide efficient methods for both identification and mitigation. New developments in machine learning methodologies have encouraging prospects for creating reliable models that can precisely identify and forecast stress levels. These methods enable us to investigate intricate linkages and patterns in multidimensional datasets, which results in more accurate stress evaluations. By comprehending the fundamental causes of stress, personalized solutions that are suited to each person's particular requirements might be created. We can provide targeted interventions geared at lowering stress and enhancing general well-being by finding patterns and correlations between different qualities and stresses. By shedding light on the variables affecting stress levels and providing direction for the creation of practical stress-reduction techniques, this study aims to make a positive impact on both research and practical applications. Our goal is to make it easier to apply evidence-based stress-reduction and resilience-building interventions by bridging the theory-practice gap.

1.4 Research Questions

1. Which characteristics are most predictive of stress levels?
2. When it comes to recognizing stress patterns, how do machine learning algorithms compare?
3. Is it possible to develop a trustworthy model for stress detection in a variety of settings?
4. How may individualized therapies be informed by stress patterns?
5. How do feature approaches affect the performance of the stress detection model?
6. Use of research results in practical stress-reduction strategies?

1.5 Expected Output

The expected output includes the construction of accurate stress detection models, insights into feature importance, and actionable recommendations for personalized stress management interventions.

Here are some basic facts I would like to mention which are related to the expected output:

- Developed Models
- Algorithm Evaluation
- Feature Importance Insights
- Pattern and Correlation Analysis
- Practical Recommendations

Developed Models

Development of strong machine learning models capable of reliably identifying and forecasting stress levels on the basis of a variety of characteristics, including behavioral, physiological, and environmental elements.

Algorithm Evaluation

An in-depth analysis of the effectiveness of several machine learning algorithms, including Support Vector Machines, Random Forests, and Gradient Boosting Machines, in identifying complex patterns in stress data.

Feature Importance Insights

Determining critical features by means of careful feature selection and engineering techniques, exposing major factors impacting stress levels.

Pattern and Correlation Analysis

Detailed examination reveals underlying trends and connections between the stressors and features found, providing insightful information to guide customized stress management strategies.

Practical Recommendations

Converting study results into doable suggestions that will help develop evidence-based plans for stress reduction and general well-being enhancement in practical settings.

1.6 Project Management and Finance

Project Management

Project Planning: I created a very detailed plan for my project. This plan included specific times when I needed to check how things were going. These checks were like important stops along the way. My supervisor helped me make sure I was on track. For example, I had checkpoints when I collected data, prepared it for analysis, built my models, and checked how well they worked. I also made sure to check in regularly to see if everything was going according to my plan. This way, I could catch any problems early and make sure I finished each part of the project on time.

Resource Management: Effective resource management means making sure I use my computers and money wisely to collect data, build models, and analyze it. I made sure to find tools and resources that would help me get things done without spending too much money. This involved carefully planning how to allocate my resources and looking for cost-effective solutions wherever possible. By being smart about resource management, I was able to accomplish my goals efficiently while staying within my budget.

Collaboration: Lines of communication were open with all relevant parties, such as funding sources, institutional partners, and research colleagues. Collaboration and support were encouraged through regular updates and feedback sessions.

Time management: Ensured timely completion of projects by closely monitoring progress and revising plans as required. Prioritized work according to my importance and urgency to complete projects efficiently within deadlines.

Finance

Budget Allocation: The main computing resources required for the research were GPUs, which were managed prudently within the allocated budget. I opted for a subscription-based GPU for my project, which amounted to 1500 BDT. Additionally, I allocated an extra 1000 BDT for data collection, as I physically gathered the data by visiting hospitals. To ensure efficient utilization of these resources, I closely monitored their usage and explored cost-saving measures where feasible. This involved optimizing code and scheduling data collection trips efficiently to minimize expenses. By carefully managing these resources, I was able to maintain financial discipline while conducting high-quality research.

Cost-Benefit Analysis: The investment in computer resources to train multiple additional models was justified by the expected savings in long-term computing expenses. The study's conclusions regarding the efficacy of several models—KNN, Gradient Boosting, SVM, Random Forest, Decision Tree, Logistic Regression, and Neural Networks—indicatively showed plausible prospective contributions to neural network architecture optimization.

Financial Implications: The conclusions about the efficacy of substitute models have significant monetary consequences. Through resource utilization optimization, the research proposes possible cost savings in computational charges for a variety of businesses. Adopting the effective models that the study found may result in significant financial gains.

1.7 Report Layout

The report contains several chapters which are Introduction, Background, Methodology, Result and Analysis, Impacts in Society and the Environment, Conclusion, and the Future Implications.

In the Chapter 1, I've provided a core understanding of Stress detection, highlighting their importance, and outlining main challenges, while also mentioned my contributions to the field. In the Chapter 2 it delves into a thorough expansion of related works, mentioning the limitations of existing methods with machine learning weights and activations, and asserting how machine learning models offer higher performance compared to conventional approaches. Entering Chapter 3, I have described in great depth the architecture and dataset that I used for my research, illuminating the complexities of the work's mechanism, and providing a foundation for the analysis that follows. The results and analysis section of Chapter 4 carefully examines the collected data and compares it to previous research to confirm the validity of the findings and the progress made in the field. A more comprehensive perspective is provided in Chapter 5, which examines the extensive advantages and effects of machine learning and offers a comparative study that highlights the importance of this work in relation to the area. As the voyage comes to an end, the last Chapter 6 opens up possibilities by outlining possible paths for more research and development in the field of machine learning.

CHAPTER-2

BACKGROUND

2.1 Preliminaries / Terminologies

In the preliminary section of this report, key terminologies include stress, machine learning, features, models, supervised learning, unsupervised learning, feature selection, evaluation metrics, cross-validation, overfitting, underfitting, hyperparameters, feature engineering, and bias-variance tradeoff.

2.2 Related Works

The landscape of stress detection research is characterized by a rich tapestry of studies exploring diverse methodologies to identify and understand stress patterns. The recognition of excess stress as a precursor to medical conditions and potentially fatal illnesses has spurred a wave of endeavors to develop cost-effective and non-invasive stress detection systems.

Pioneering studies, such as the work by Non-invasive portable devices using body temperature, heart rate, and galvanic skin response (GSR) were first shown by Pascual et al. [1]. portable sensors played a crucial role, enabling real-time monitoring of stress levels in scenarios ranging from oral presentations to examinations.

Expanding the spectrum, researchers have explored remote data acquisition using Global System for Mobile Communications (GSM) and Global Positioning System (GPS) [7]. Additionally, the PhysioNet database [8][9][10] served as a valuable resource for accessing datasets of physiological signals like ECG, GSR, and respiration rate.

Incorporating advanced metrics, studies ventured into electroencephalogram (EEG)-based stress detection [11], employing channel electrodes to obtain physiological data signals from volunteer health. Furthermore, the exploration of biological data, such as respiration, GSR, heart rate, and electromyogram (EMG), for stress assessment in contexts like driving [13] demonstrated the versatility of stress detection parameters.

Machine learning algorithms, a cornerstone in stress detection, have been rigorously applied across studies. In a comparison of techniques such as Support Vector Machine (SVM), Linear Discriminant Analysis (LDA), Ensemble, K-nearest neighbor (KNN), and Decision Tree (DT), Hasanbasic et al. [1] demonstrated the superiority of SVM with an amazing 91% accuracy. Diversified algorithmic comparisons [3][10] underscored the significance of tailored approaches, with SVM, Logistic Regression (LR), and Artificial Neural Network (ANN) exhibiting high accuracies.

Despite these advancements, previous studies have yet to explore the combined potential of GSR, systolic BP, diastolic BP, and HR as a stress detection ensemble. Moreover, gender differences, a significant factor in stress variations, have been underexplored as a feature in stress detection models.

This study aims to bridge these gaps by refining machine learning models to achieve higher accuracy and reliability. By combining specific physiological parameters and considering gender as a feature, the research endeavors to propel the field of stress detection towards more nuanced and precise methodologies.

TABLE 2.2.1: COMPARISON BETWEEN EXISTING WORKS

Serial No	Study	Physiological Parameters Monitored	Machine Learning Algorithm(s) Used	Accuracy (%)
1	Pascual et al. [1]	GSR, Body, Temperature, Heart Rate	SVM, LDA, Ensemble, KNN, DT	91
2	M. A. Jiménez- limas, C. A. Ramírez-fuentes, B. Tovar-corona, and L. I. Garay-jiménez, [8] [9] [10]	Respiratory rate, GSR, and ECG signal	KNN, Linear SVM, Gaussian SVM, Naive Bayes, RF, LR	81.38
3	Betti et al. [2]	Heart rate, electrodermal, and electroencephalography signals	SVM	86
4	Zhang [15]	ECG	SVM	93.75
5	Vanitha & Suresh [16]	HRV from ECG	SVM	91
6	O. Sakri, C. Godin, G. Vila, E. Labyt, S. Charbonnier, and A. Campagne, [3] [13]	Galvanic skin response (GSR), body temperature, and heart rate.	SVM, DT, Naive Bayes	85
7	This Study- Predection of stress levels based on physical activity using machine learning	Gender, Age, Sleep-Duration, Sleep-Hour, Physical-Activity- Level, Stress- Level, BMI Category, Systolic-Pressure, Diastolic-Pressure, Heart- Rate Daily-Steps	SVM, KNN, LR, RF, Neural Network, Decision tree, Gradient Boosting Machine (GBM)	95

2.3 Comparative Analysis and Summary

An Additional Paperwork:

Dataset: Consists of some characteristics (gender, galvanic skin response, systolic and diastolic blood pressure, and heart rate) Individuals from a range of age groups provided data, which made the sample for the analysis diversified. These attributes include information on physiological measures as well as personality characteristics, allowing a thorough investigation of stress levels across various age groups.

Methodology: Grid search was used for optimization while machine learning algorithms such as Support Vector Machine, K-Nearest Neighbor, and Logistic Regression were used.

Findings: When it came to stress level classification, SVM performed the best, with an optimum accuracy score of 93%.

My Paperwork

Feature Extension: I increased the number of features to eleven in an attempt to capture an increased number of physiological and contextual variables that affect stress.

Model Diversity: I made use of a greater variety of models, such as SVM, KNN, Random Forest, Gradient Boosting, Decision Trees, and Neural Networks (MLP).

Findings: Neural Network (MLP) showed somewhat lower accuracy (85%), followed by SVM (84%), KNN (81%), and Logistic Regression (71%), while Random Forest, Gradient Boosting, and Decision Tree models reached similar high accuracies (95%).

Gap Analysis:

Feature Set Expansion: Previous studies have mostly focused on a few physical characteristics to detect stress. My study aims to close this gap by expanding the feature set to include more variables.

Model Performance Diversity: To classify stress, previous studies have frequently depended on a small number of machine learning models, with a focus on Support Vector Machine (SVM). By contrast, a wide range of models—including some that haven't been studied before, such as Random Forest, Decision tree and Gradient Boosting—are included in my research.

Practical Implications: Although previous studies have highlighted the significance of taking into consideration diverse models for stress detection systems, there is still a deficiency in converting these discoveries into useful implementations. Research on the viability and scalability of implementing these models in practical contexts is required, taking into consideration problems with data collecting and processing capacity. Furthermore, there is a lack of knowledge regarding the particular use cases and limitations that affect which model is best suited for a given application. To close the gap between research findings and real-world applications, future studies may concentrate on creating frameworks or standards for model selection based on practical factors.

Summary

Both studies make contributions to the field of machine learning-based stress detection; however, my paper improves the feature set and assesses several models, while the other paper concentrates on improving SVM for classification based on a restricted feature set. My study provides a more advanced knowledge of stress detection and highlights the significance of taking into consideration many elements when constructing effective classification systems by integrating extra features and investigating diverse model techniques. The two publications' comparative research shows how feature diversity and expansion may improve the accuracy and reliability of stress detection systems.

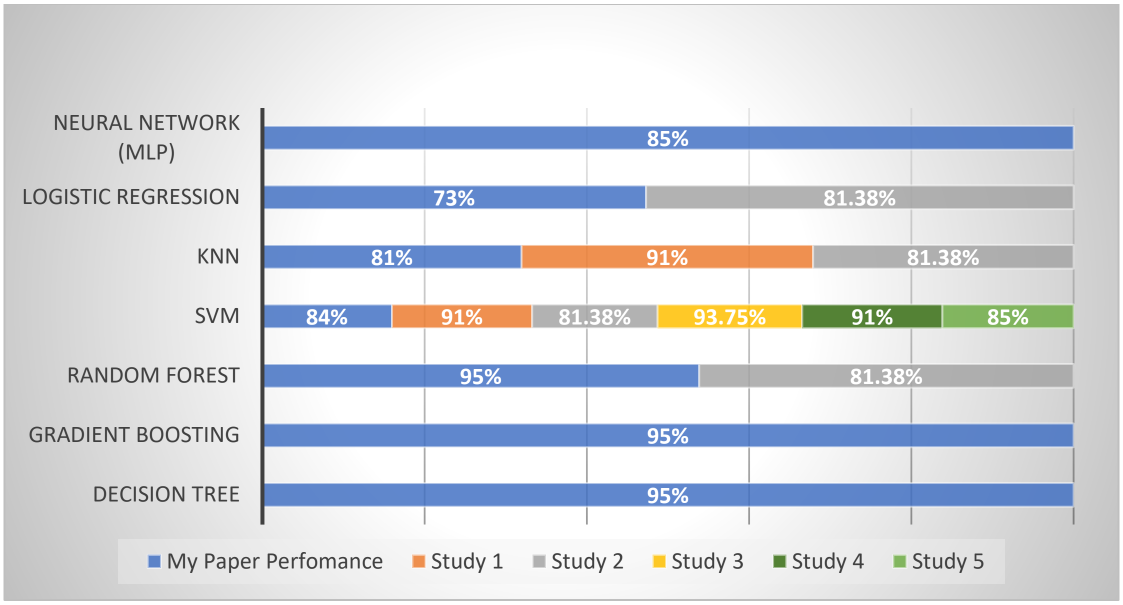

Figure 2.3.1: Comparison of Stress Level Detection Classifier Accuracies Across Studies

This figure highlights the efficacy of particular algorithms in achieving high predicted accuracy by displaying the disparate accuracies of various machine learning classifiers in stress level detection across multiple studies.

2.4 Scope of the Problem

The field of stress detection by machine learning is extensive and involves many elements. To start, it is important to recognize that stress is a result of complex interactions between several different factors rather than a single, separate thing. These variables include psychological, environmental, and social components in addition to physiological signs like blood pressure and heart rate. The scope of the study is extensive to include individuals from a range of age groupings, genders, occupations, and socioeconomic backgrounds in recognition of the varied demographics that are impacted by stress. Furthermore, stress is an evolving state that changes with time and in reaction to various stressors. As a result, the scope includes techniques for real-time monitoring and temporal data gathering to record both acute stress events and long-lasting stress patterns.

Recognizing the moral problems and data privacy concerns that arise from the collecting and analysis of stress-related data is an essential component of this focus. To protect participants' right to privacy, the framework has clauses for getting informed consent, anonymizing sensitive data, and following legal requirements. Furthermore, the scope encompasses a range of application contexts, such as educational institutions, workplace wellness initiatives, healthcare settings, and personal wellness management systems. The intention of focusing on these areas is to offer flexible and easily integrable solutions for use with wearables, mobile apps, and web platforms, among other systems and technologies. Essentially, researchers and professionals can create customized techniques that meet the complex nature of stress while maintaining morality and usefulness by thoroughly outlining the problem's depth.

To ensure the relevance and applicability of my paper to the intended users, the stress detection system, for example, highly educated adults aged 18 to 70 years, was developed with patients. Regardless of whether the stress detection system is being used as a personal stress management tool, in corporate wellness efforts, or in healthcare settings, the intended use cases have influenced the system's design and implementation. This comprehensive approach helps formulate realistic goals, design appropriate protocols, and interpret results effectively, all of which advance stress detection and management.

2.5 Challenges

To successfully deal with the difficulties associated with stress detection using machine learning, an integrated approach that takes into consideration the complex structure of the procedure is needed.

The following are some key challenges:

- Subjectivity and Variability
- Feature Selection and Data Quality
- Overfitting and Generalization
- Validation and Clinical Adoption
- Interdisciplinary Collaboration

Subjectivity and Variability

The subjectivity and variation of stress experiences are one of the main difficulties. Since stress appears differently in individuals and situations, creating a detection model that works for everyone is difficult. Subjective elements like perception and coping techniques also add to the process complexity.

Feature Selection and Data Quality

For the model to work well, appropriate characteristics that properly capture stress signals must be found. However, choosing relevant characteristics from different datasets and maintaining data dependability and quality can be difficult, particularly when working with self-reported or sensor-based data.

Overfitting and Generalization

Overfitting is a condition in which machine learning models that have been trained on certain datasets perform well on training data but badly on unseen data. Reliable stress detection across a range of people and environments demands finding a compromise between model complexity and generalization capacity.

Validation and Clinical Adoption

There are many challenges to overcome for stress detection models to be accepted by end users and healthcare providers as well as validated in real-world scenarios. Thorough validation studies, evidence-based decision-making, and integration with current healthcare workflows are necessary for clinical adoption.

Interdisciplinary Collaboration

The complex problem of stress detection requires cooperation between computer science, engineering, psychology, and medicine. Creating machine learning models that are both successful and comprehensible requires bridging the gap between knowledge of the topic and technological proficiency.

CHAPTER 3

RESEARCH METHODOLOGY

3.1 Research Subject and Instrumentation

Research Subject

The primary objective of this study is to create a reliable system that can accurately identify stress levels. I have used 6 models out of which 3 models are getting high assurance of 95%, they are Machine Learning Random Forest, Gradient Boosting, Decision Tree. Using a variety of models with 95% confidence levels, my goal was to create an all-encompassing framework that could precisely measure stress in various populations. The study involved the integration of various characteristics derived from the stress levels of both patients and normal individuals experiencing stress-related conditions. My objective is to enhance the efficacy and dependability of stress detection by the utilization of diverse model sets and features. This will ultimately lead to the development of more targeted stress management techniques and better health care treatments.

- Here I have collected data of 393 people whose age is 20 to 70 years.
- I have created some models that give better results than other papers and model.
- I have collected the data in 2 ways 1 from medical patients and other from common people.

Instrumentation

In this field of study, “instrumentation” refers to a wide variety of instruments, devices, and techniques that are employed to gather information on factors associated with stress and assist in the creation of machine learning models. Researchers aim to develop stable and dependable methods for stress assessment that can be used in a variety of contexts, such as personal stress management, workplace wellness, and healthcare, by integrating these instruments. These are:

- **Google Colab’s**

Platform for data processing and analysis: Made effective use of Google Colab's computational capacity and teamwork tools for handling and analyzing data.

Environment for Model Development: Several machine learning models for stress detection were developed and trained at Google Colab’s using the Python programming environment.

- **Diverse models**

Selection Criteria: I have chosen several models known for their effectiveness in classification tasks: Support Vector Machine (SVM), Random forest, Gradient Boosting, Neural network, Decision tree, Logistic regression.

Assurance Level: Models with a minimum assurance level of 95% were given priority, guaranteeing high levels of confidence in the results of stress detection.

- **Entire Dataset**

Data Collection: Collected a large dataset with a variety of characteristics related to stress levels. Physiological measures, self-reported symptoms, lifestyle factors, etc. are examples of features. For a balanced dataset, it was made sure that both patients with stress-related conditions and people with normal stress levels were represented.

- **Plots and Graphs**

Visualization Tools: To visually examine the dataset and model performance, a variety of Python graphing and plotting packages, such as Matplotlib and Seaborn, were used. Gained insights into stress patterns and model efficacy by creating visual representations of data distributions, correlations, and model evaluation measures.

- **Metrics for Assurance**

Evaluation Criteria: To assess the quality of reliability and confidence in each model's predictions, assurance metrics were applied. Area under the ROC curve (AUC), recall, precision, accuracy and F1-score are some examples of metrics. To ensure that the models are suitable for stress detection in practical applications, we set a minimum threshold of 95% certainty.

The research aims to develop a trustworthy stress detection system that can precisely determine stress levels in a variety of populations by utilizing these equipment components.

3.2 Data Collection Procedure/Dataset Utilized

Two main methods were used in this study to collect data: patient medical files and publicly accessible datasets from websites such as Kaggle that show healthy individuals. The goal of integrating these datasets was to record various stress-related variables in different age categories.

Data Collection Procedure

Medical Dataset Collection:

Collaborated with healthcare institutions, clinics and medical professionals to access medical records and data from patients experiencing stress-related conditions. Obtained consent from patients and ensured compliance with ethical guidelines regarding data privacy and confidentiality. Relevant information was extracted from medical records, including diagnostic evaluation, medical history, physiological measurements, and self-reported symptoms.

Obtaining the Kaggle dataset:

Publicly accessible datasets from reliable platforms like Kaggle were employed, focusing on information about mental health, lifestyle factors, and socio-demographic features. In order to represent a varied and representative sample of the general public, including people with normal stress levels, datasets have been evaluated and chosen.

Data Integration:

Medical and Kaggle datasets are merged to create a unified dataset combining patient and general population information. Variable definitions, formats, and units to ensure consistency across integrated datasets. By conducting data cleaning procedures to address missing values, outliers, and inconsistencies, dataset quality and reliability are enhanced.

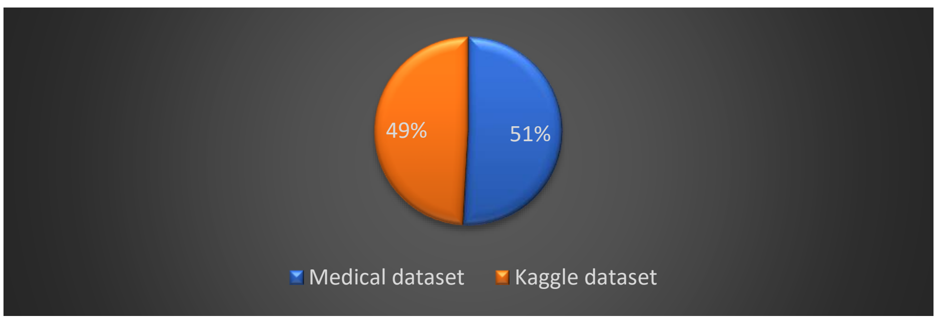

Figure 3.2.1: Ratio of Medical Dataset & Kaggle Dataset

Here in the graph, I can see the ratio of Medical & Kaggle Dataset. I have collected both Medical & Kaggle Dataset to train my model with both types of data so that the model gives more accurate results.

Dataset Utilized

The dataset utilized in this research is a merged compilation of medical data from patients and publicly available datasets representing normal individuals. Key features included in the dataset are as follows:

Medical Dataset (Patient Data):

- **Diagnostic assessments:** Psychological evaluations, stress-related disorder diagnoses.
- **Treatment histories:** Medications, therapies, interventions.
- **Physiological measurements:** Heart rate variability, blood pressure, cortisol levels.
- **Self-reported symptoms:** Perceived stress levels, anxiety symptoms, sleep disturbances.
- **Institution:** Rajshahi Central Hospital, Laxmipur, Rajshahi, Bangladesh.

In collaboration with Rajshahi Central Hospital in Laxmipur, Rajshahi, Bangladesh, these data were collected. The hospital made patient medical records accessible, which made it possible to gather in-depth data about stress and its contributing variables.

Kaggle Dataset (Normal Population Data):

Socio-demographic characteristics: Age, gender, education level, employment status.

Lifestyle factors: Exercise frequency, dietary habits, social support networks.

Mental health assessments: Anxiety scores, depression indicators, well-being measures.

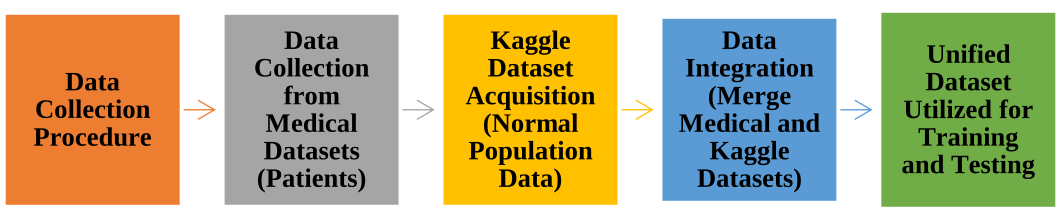

Figure 3.2.2: Visual representation of the sequential steps

This figure depicts the sequential steps involved in the stress detection process, illustrating the workflow from data collection to the application of machine learning algorithms.

After creating organized steps, I got to work. The first stage was preparing my dataset, which is an important step in data manipulation to improve model performance. Because noisy data comprised a significant component of the dataset, pretreatment was considered necessary to provide a smooth model construction procedure. The following is a summary of the steps used for dataset preprocessing.

Sample of Dataset:

TABLE 3.2.1: SAMPLE OF DIVERSE DATASET

	A	B	C	D	E	F	G	H	I	J	K
1	Gender	Age	Sleep_Duration	Sleep_Hour	Physical_Activity_Level	Stress_Level	BMI Category	Systolic_Pressure	Diastolic_Pressure	Heart_Rate	Daily_Steps
2	Male	22	7	7	70	4	Normal	120	80	74	2000
3	Female	25	7	9	50	3	Normal	90	65	78	1200
4	Male	26	7	6	40	6	Normal	110	80	72	1500
5	Male	27	6.1	6	42	6	Overweight	126	83	77	4200
6	Male	28	6.2	6	60	8	Normal	125	80	75	10000
7	Male	28	6.2	6	60	8	Normal	125	80	75	10000
8	Male	28	5.9	4	30	8	Obese	140	90	85	3000
9	Male	28	5.9	4	30	8	Obese	140	90	85	3000
10	Male	28	5.9	4	30	8	Obese	140	90	85	3000
11	Male	28	8	9	60	4	Normal	120	80	72	2100
12	Male	29	6.3	6	40	7	Obese	140	90	82	3500
13	Male	29	7.8	7	75	6	Normal	120	80	70	8000
14	Male	29	7.8	7	75	6	Normal	120	80	70	8000
15	Male	29	7.8	7	75	6	Normal	120	80	70	8000
16	Male	29	6.1	6	30	8	Normal	120	80	70	8000
17	Male	29	7.8	7	75	6	Normal	120	80	70	8000
18	Male	29	6.1	6	30	8	Normal	120	80	70	8000
19	Male	29	6	6	30	8	Normal	120	80	70	8000
20	Male	29	6	6	30	8	Normal	120	80	70	8000
21	Male	29	6	6	30	8	Normal	120	80	70	8000
22	Female	29	6.5	5	40	7	Normal Weig	132	87	80	4000

3.3 Statistical Analysis

There are some statistics followed as-

- Descriptive Statistics
- Correlation Analysis
- Hypothesis Testing
- Regression Analysis

Descriptive statistics: These statistics provided an overview of the distribution of stress-related variables, enabling researchers to identify trends and outliers.

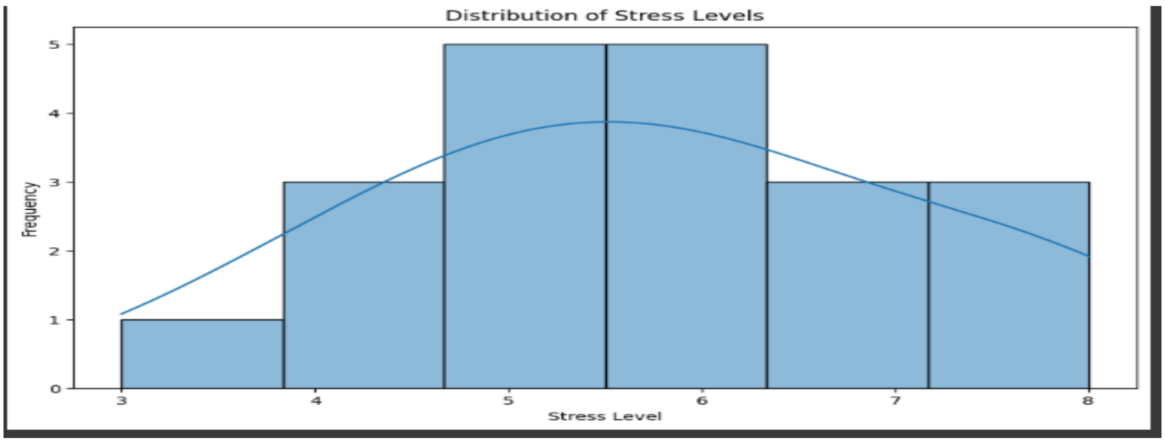

Figure 3.3.1: Distribution of Stress Levels

This figure the distribution of stress levels among participants is shown in this image, which provides light on the frequency of various stress levels in the population under study.

Correlation Analysis: To find out the connections between the various variables in the dataset, a correlation analysis was done. To measure the direction and strength of linear correlations between continuous data, Pearson correlation coefficients were computed. When analyzing non-linear correlations or ordinal variables, Pearson correlation coefficients were applied.

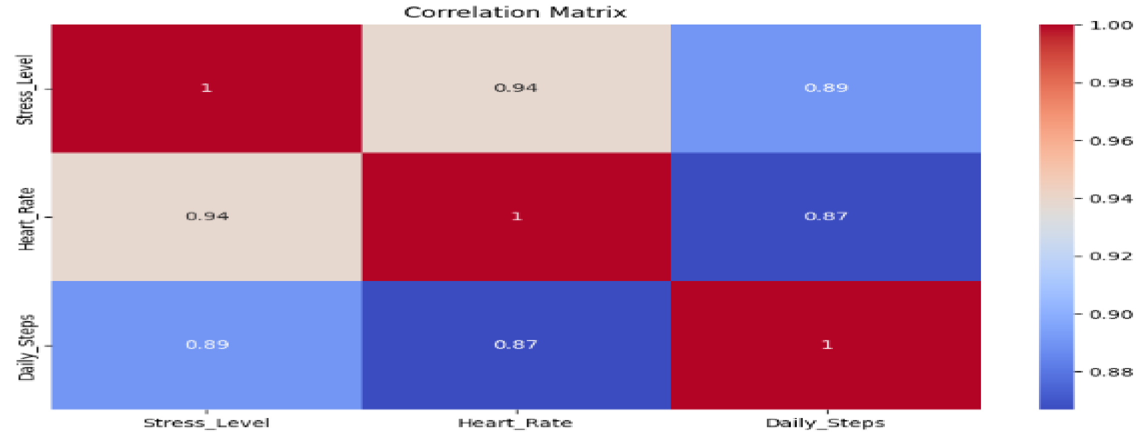

Figure 3.3.2: Correlation Matrix

This figure displays the correlation matrix, highlighting the relationships between various physiological and behavioral features used in the stress detection study.

Hypothesis Testing: To evaluate the importance of connections or distinctions between groups within the dataset, hypothesis testing was used. When comparing means between two groups—for example, the stress levels of sick and normal people—t-tests were employed. When comparing more than two groups, analysis of variance analyses were used, and after the fact testing was carried out to find particular group differences.

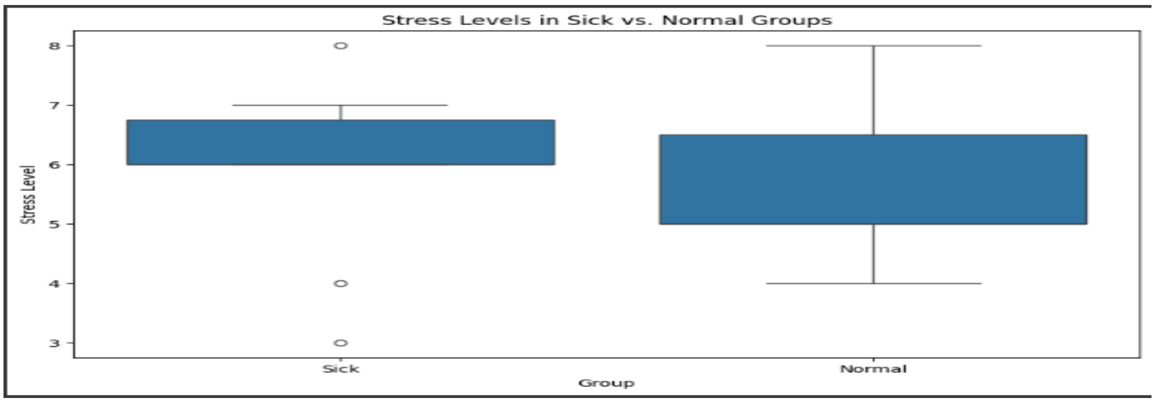

Figure 3.3.3: Sick Vs Normal Group

This figure contrasts physiological and behavioral metrics between groups experiencing elevated stress levels ("Sick") and those with normal stress levels, providing insights into distinguishing features.

Regression Analysis: The associations between independent variables and stress levels were modeled using regression analysis approaches, such as logistic regression. While logistic regression models were utilized for binary classification tasks, linear regression models were utilized to predict continuous stress levels based on predictor factors.

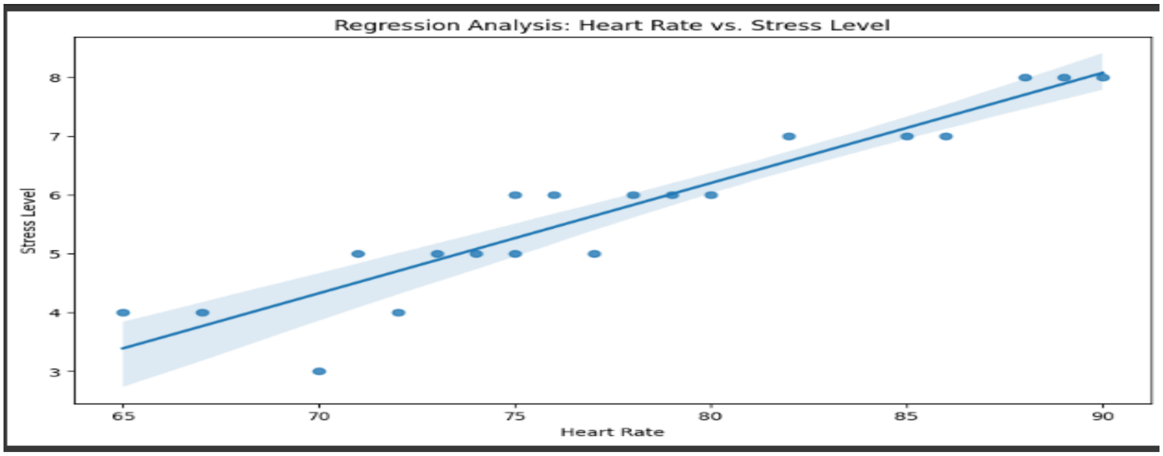

Figure 3.3.4: Regression Analysis

This figure presents the results of regression analysis, exploring the relationship between stress levels and various predictors, highlighting significant factors influencing stress outcomes.

Interpretation of the Results

- The statistical analysis's findings shed light on the variables—such as physiological parameters, lifestyle choices, and demographic traits—that affect stress levels.
- Prioritizing pertinent features for inclusion in stress detection models was made easier by significant correlations and linkages found through statistical testing.
- The results of regression analysis made it easier to identify predictive variables and create models that could reliably predict stress levels.

Implications

- The results of the statistical analysis formed the basis for the model's development, directing the choice of features and influencing the architecture of the model.

- The study provided insights that assisted in the creation of focused interventions and individualized stress management techniques by providing a greater knowledge of the complex interactions between the many stressors.

3.4 Proposed Methodology/Applied Mechanism

Proposed Methodology

Data Collection: Gathering diverse physiological and behavioral data, including heart rate, sleep patterns, and physical activity levels.

Data Preparation: Cleaning, preprocessing, and feature engineering to ensure data quality and relevance for modeling.

Model Training: Utilizing machine learning algorithms such as SVM, KNN, and LR to build predictive models based on the prepared dataset.

Visualization and Interpretation: Visualizing model outputs and interpreting patterns and correlations in the data to derive meaningful insights.

Model Evaluation: Assessing model performance through metrics like accuracy, precision, and recall to validate its effectiveness in stress level prediction.

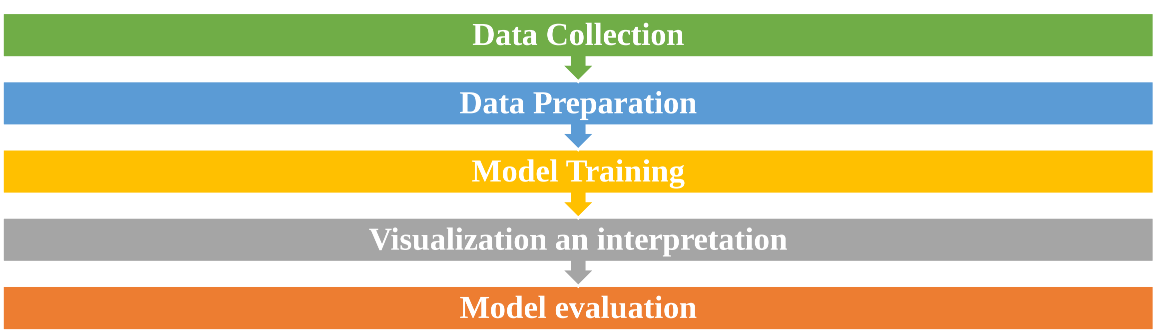

Figure 3.4.1: Visualized Proposed Methodology

So, Now Here I show some visual representation of pair plots, bar graphs, box plots, and histograms.

Pair plots:

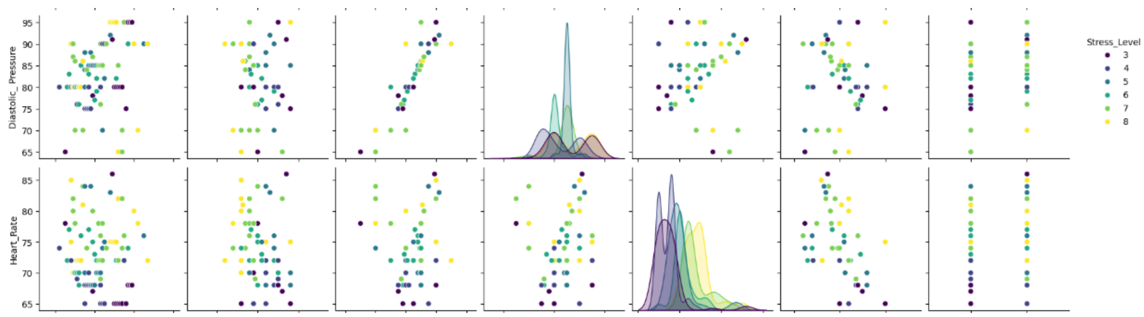

Figure 3.4.2: Pair plot of Numerical Features

This figure displays a pairwise scatter plot matrix of numerical features, illustrating the relationships and distributions between variables used in stress level analysis.

Bar graphs:

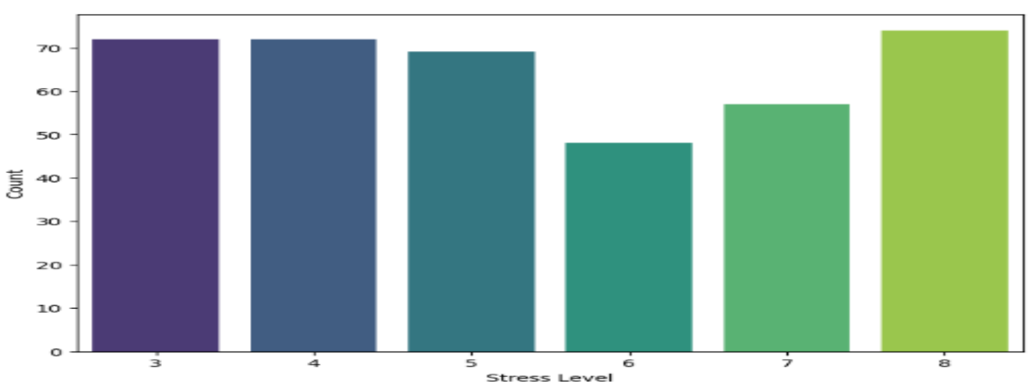

Figure 3.4.3: Stress Level Distribution

This figure visualizes the distribution of stress levels across the study population, providing insights into the prevalence and intensity of stress experienced by participants.

Boxplot:

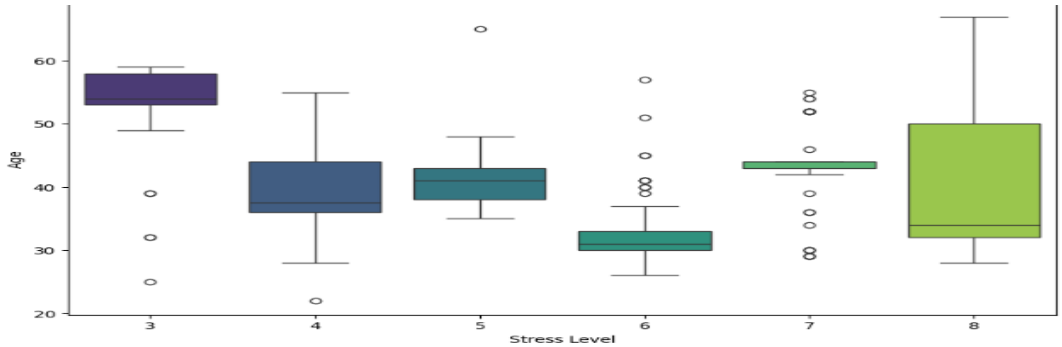

Figure 3.4.4: Age Distribution by Stress Level

This figure illustrates how age correlates with different levels of stress among participants, highlighting potential age-related patterns in stress experiences.

Histogram:

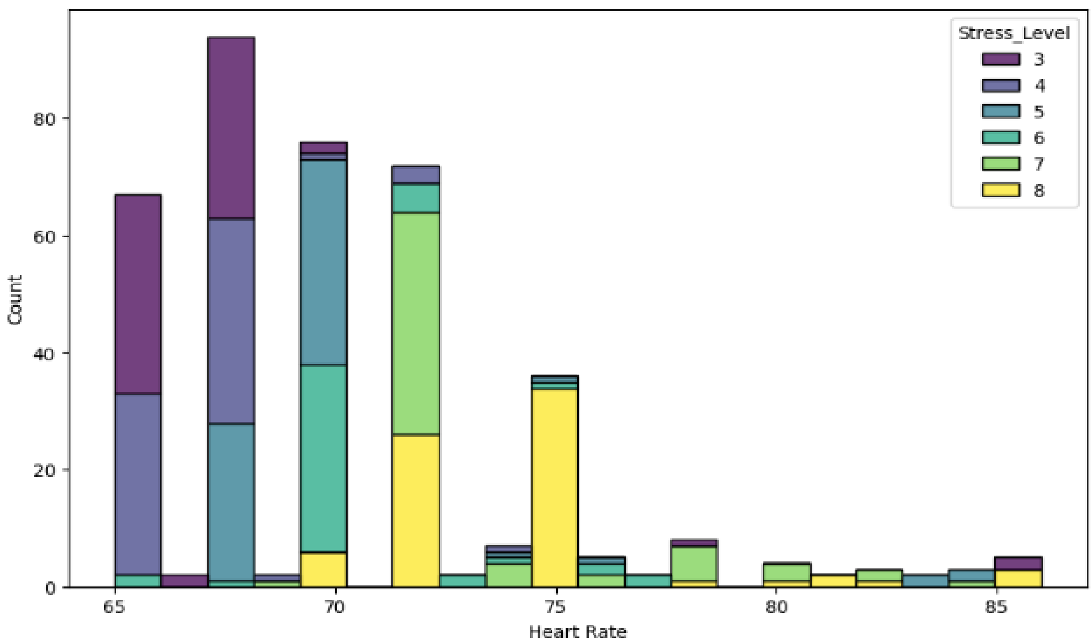

Figure 3.4.5: Heart Rate Distribution by Stress Level

This figure depicts the distribution of heart rates across varying levels of stress, providing insights into how physiological metrics correlate with stress intensity among participants.

Model Training and Evaluation

Dataset Splitting: To make model training and evaluation easier, the dataset is divided into training and testing sets. Usually, 80/20 or 70-30 splits are employed, with training receiving the majority of the data.

Model Initialization: Default hyperparameters are used to start a number of classifiers, such as Random Forest, Gradient Boosting, Decision Tree, Logistic Regression, and Support Vector Machine (SVM).

Model Training: The training dataset is used to train each classifier. In order to identify the underlying patterns and correlations, the model is fitted to the training set of data during the training process.

Evaluation of the Model: The testing dataset is used to assess the trained classifiers. Performance measures including recall, accuracy, precision, F1-score, and confusion matrix are calculated to evaluate how well each model detects stress levels.

3.5 Implementation Requirements

To successfully apply the suggested methodology for stress level detection, the following requirements need to be fulfilled:

Python settings: Installed Python on my PC. Made sure I installed all required libraries including scikit-learn, matplotlib, seaborn, pandas and numpy.

Dataset: Take a dataset with specific features related to stress level. To preprocess and accurately represent the dataset, including encoding missing values and categorical variables.

Data Preprocessing Tools: Tools and methods using data from priors, such as immediate encoding for categorical variables, scaling numerical features and, if necessary, dealing with imbalanced data.

Machine Learning Algorithms: Used a variety of machine learning techniques, such as support vector machine (SVM), gradient boosting, random forest, logistic regression, and decision trees to detect stress levels. Make sure I'm familiar with the hyperparameters of the algorithm.

Metrics for Model Evaluation: I have established appropriate evaluation systems to measure the performance of trained models. F1-score, precision, area under recall are examples of common metrics.

Tools for Data Visualization: To show the dataset's distributions, feature correlations, and model evaluation outcomes, use data visualization tools like Matplotlib and Seaborn.

CHAPTER-4:
EXPERIMENTAL RESULT AND DISCUSSION

4.1 Experimental Setup

The experimental setup for the study included the following components and techniques:

Data Acquisition: I took a diverse dataset with features related to stress levels. We included a wide range of demographics, lifestyle characteristics, physiological measures, and self-reported stress-related symptoms in this dataset.

Data Preprocessing: To guarantee data quality and machine learning algorithm consistency, the dataset is preprocessed. It addresses any class imbalances that may exist, scales numerical features, encodes categorical variables, and handles missing data.

Feature Selection: To find the most informative features for stress level detection, perform feature selection. To choose pertinent characteristics, apply methods like feature importance from tree-based models, correlation analysis, or domain expertise.

Model Selection: I selected a range of machine learning models that are known to do well in classification tasks. Examples are neural networks, tree-based models, linear models, and ensemble methods. We considered the performance, interpretability and computational complexity of each model.

Cross-Validation: We used cross-validation techniques, such as k-fold cross-validation, to assess the generalization performance of the model. To obtain reliable performance estimates, we made sure the models were evaluated using several data folds.

Performance Metrics: Established appropriate performance measures to evaluate the effectiveness of the models in detecting stress levels. F1-score, precision, accuracy, area under recall are examples of common metrics.

Experiment Execution: With the preprocessed dataset, each model was trained and evaluated using the experimental setup described above. I kept track of each model's performance metrics so I could evaluate how well it detects stress levels.

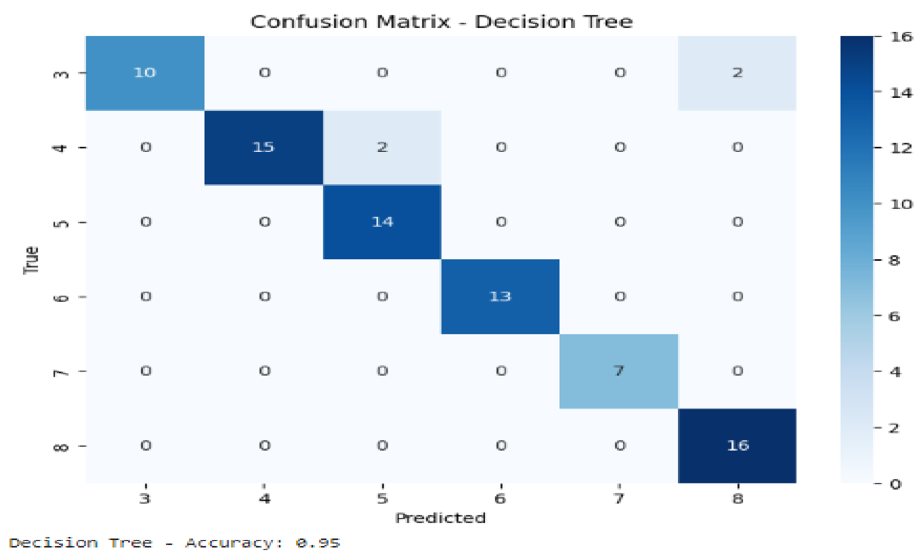

Figure 4.1.1: Training Decision Tree

In this Figure 4.1.1: Training Decision Tree, the classifier's training procedure for identifying stress levels is shown in this graphic. The Decision Tree model's remarkable 95% accuracy rate shows how well it can differentiate between various stress levels using the features that were supplied.

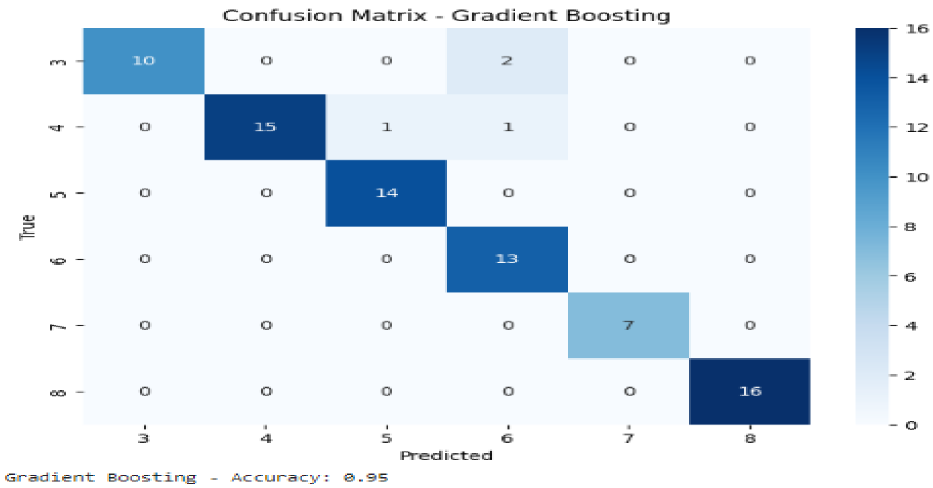

Figure 4.1.2: Training Gradient Boosting

In this Figure 4.1.2: Training Gradient Boosting, the training procedure of a Gradient Boosting classifier for stress level detection is shown in this figure. The Gradient Boosting model performed admirably, with an accuracy of 95%, demonstrating its strong ability to classify stress levels appropriately based on the given features.

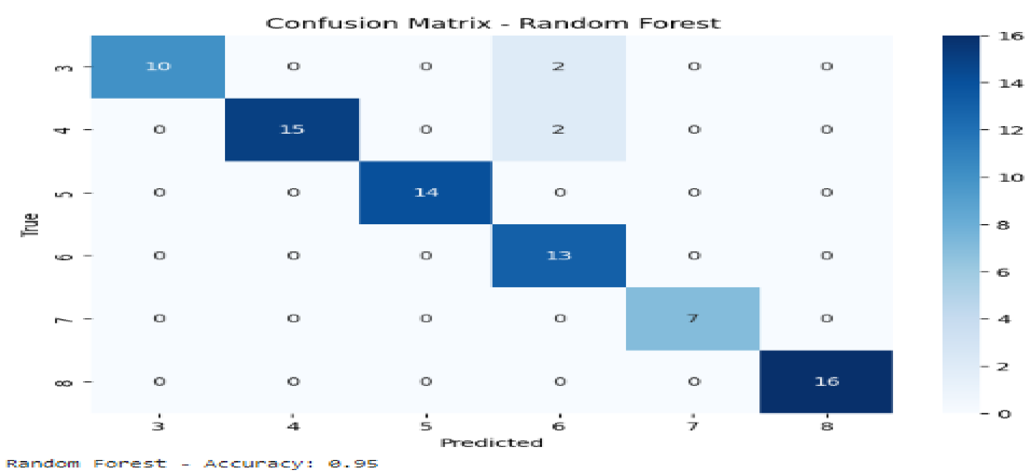

Figure 4.1.3: Training Random Forest

In this Figure 4.1.3: Training Random Forest, the Random Forest classifier's training procedure for identifying stress levels is shown in this graphic. The Random Forest model was found to be effective in accurately classifying stress levels by achieving a 95% accuracy rate based on the features provided.

4.2 Experimental Results & Analysis

The outcomes of the studies are presented in the experimental data and analysis section, which also offers insights into the connections between physiological measures, stress levels, and demographic information. It includes the following essential elements:

Model Performance: I assessed each machine learning model's effectiveness on the test dataset, using appropriate performance metrics. I compared measures under recall, precision, accuracy and F1-score to evaluate the performance of stress level detection.

Comparative Analysis: To choose the best method for stress level detection, I compared the performance of several models and examined the performance of each model in terms of computational efficiency, robustness against noise, and sensitivity to different stress levels.

Limitations and Challenges: I considered variables including class imbalance, computing resource limitations, feature engineering complexity, and data quality to talk about any restrictions or difficulties during the trial.

Practical Implications: I highlighted how improved stress detection systems can help individuals, companies, or healthcare professionals better manage stress by talking about the implications of experimental results for real-world applications.

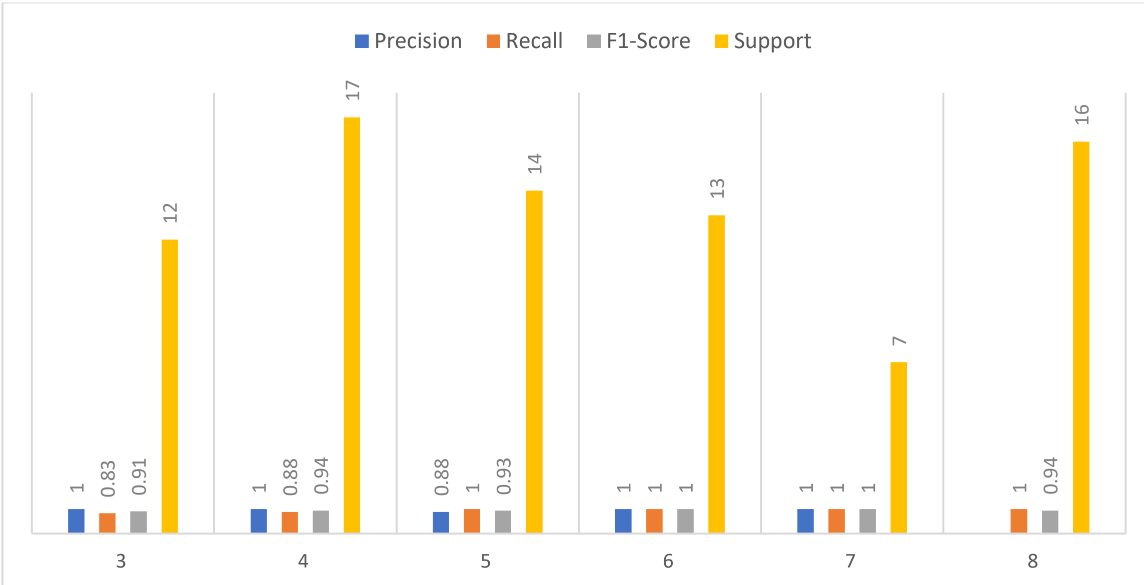

Figure 4.2.1: Visualization of Decision Tree Classification Metrics

Precision: Out of all the examples predicted as belonging to a particular class, this statistic shows the percentage of accurately predicted instances of that class. For instance, the classifier achieved 100% precision for Stress Level 3, indicating that all cases predicted to be Stress Level 3 were accurate.

Recall: Recall, sometimes referred to as sensitivity, is the percentage of accurately predicted instances of a given class out of all instances of that class that actually occur. For example, the classifier successfully identified all real instances of Stress Level 5 with a 100% recall rate for Stress Level 5.

F1-Score: This score achieves a balance between recall and precision by taking the balanced mean of both variables. It is a metric for assessing the accuracy of a model; values nearer 1 indicate superior performance.

Support: In the test dataset, support is the actual number of instances of each class. It indicates the number of cases that are available for assessment for each class. In the end, it provides a view of the metrics' reliability for each class by displaying the distribution of each class in the dataset.

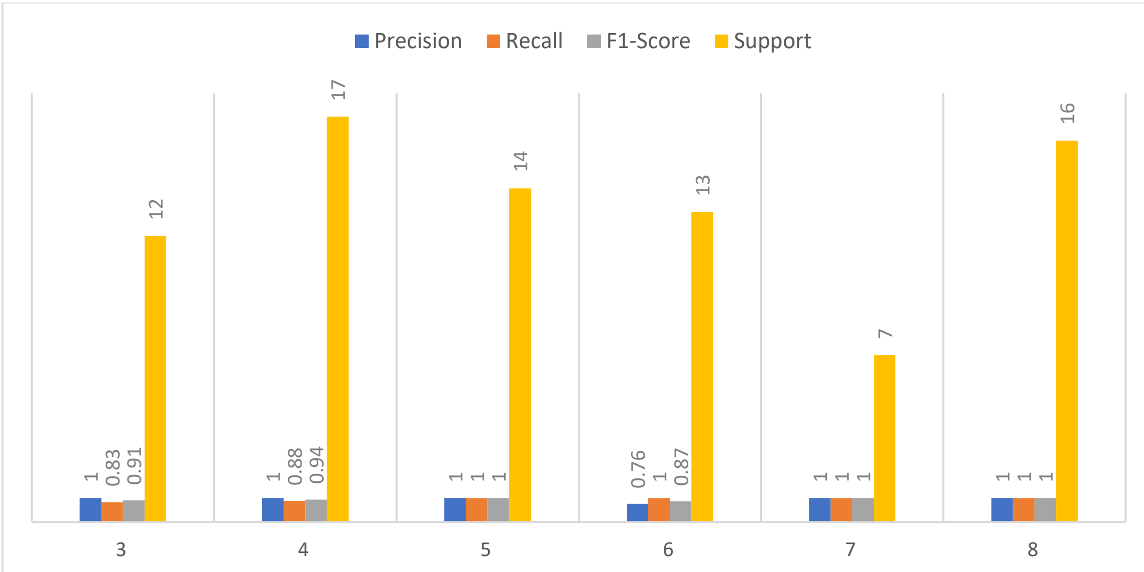

Figure 4.2.2: Visualization of Random Forest Classification Metrics

Precision: Precision is the percentage of accurately predicted instances of a given class out of all instances predicted to belong to that class. For example, the Random Forest classifier for Stress Level 3 achieved 100% precision, meaning that all the instances that were predicted to be Stress Level 3 were accurate.

Recall: Recall, sometimes referred to as sensitivity, is the percentage of accurately predicted instances of a given class out of all instances of that class that actually occur. For instance, the classifier successfully identified all true occurrences of Stress Level 6 with a 100% recall rate for Stress Level 6.

F1-Score: This score achieves a balance between recall and precision by taking the balanced mean of both variables. It is a metric for assessing the accuracy of a model; values nearer 1 indicate superior performance. For instance, the Stress Level 5 F1-Score is 1.00, which denotes excellent recall and precision.

Support: In the test dataset, support is the actual number of instances of each class. It indicates the number of cases that are available for assessment for each class. For example, Stress Level 4 has a support of 17, which indicates that the test dataset contains 17 instances of Stress Level 4.

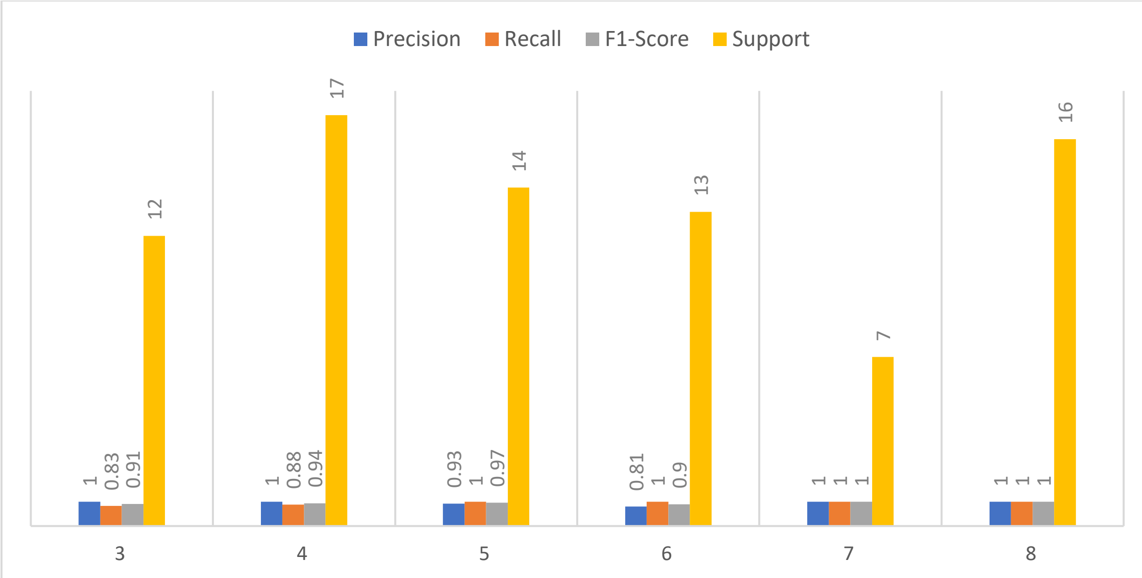

Figure 4.2.3: Visualization of Gradient Boosting Classification of Metrics

Precision: Shows the percentage of all instances accurately predicted to belong to a certain class out of all instances predicted to belong to that class. For example, Stress Level 3 attained 100% precision, meaning that every case that was projected to be Stress Level 3 was accurate.

Recall: Indicates the percentage of accurately predicted instances of a given class out of all instances of that class that actually occur. Stress Level 5, for instance, had a 100% recall rate, which means that every true occurrence of Stress Level 5 was recognized.

F1-Score: The F1-score finds a compromise between recall and precision by representing the inverse mean of both measurements. Greater values indicate improved model functionality.

Support: Provides information about the class distribution of the dataset by indicating the actual number of instances of each stress level in the test dataset.

Accuracy and Loss curve:

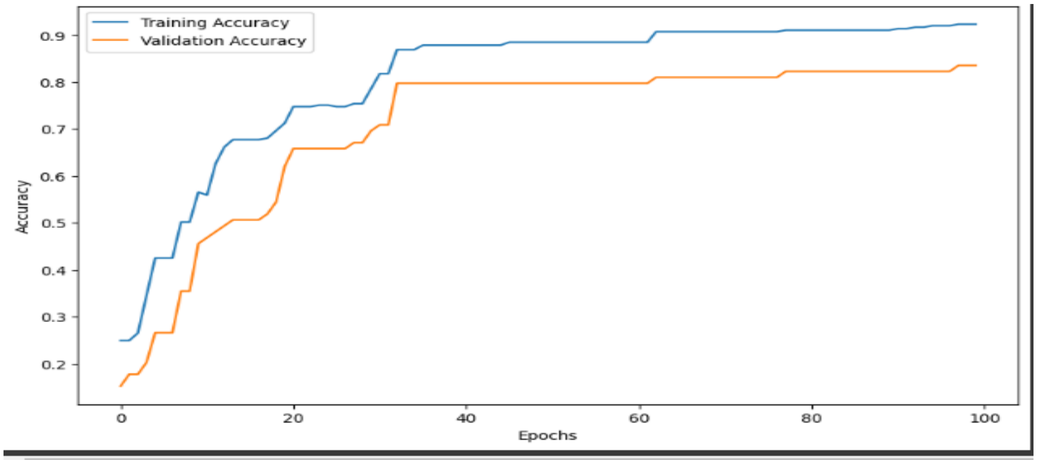

Figure 4.2.4: Accuracy Curve

This figure the accuracy curve for the MLPClassifier illustrates how the training and validation accuracies evolve over epochs. This curve helps to identify the model's learning progress, with improvements in accuracy indicating better generalization and learning efficiency.

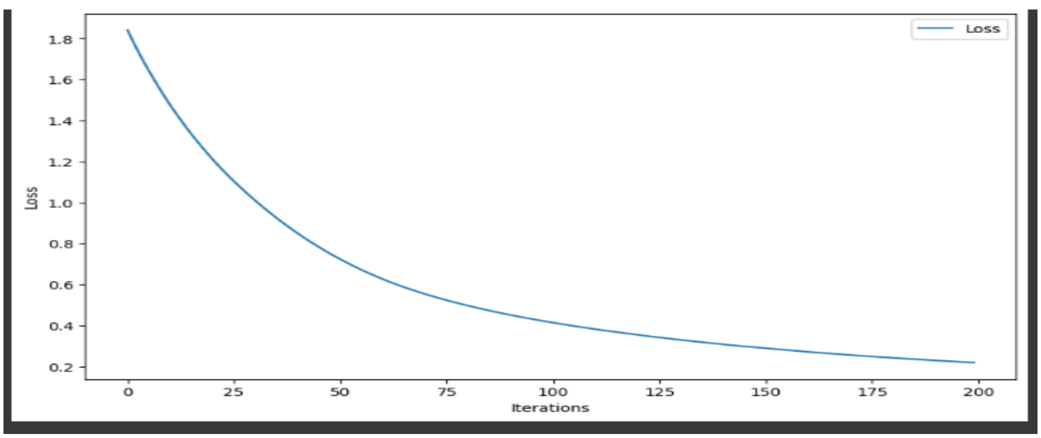

Figure 4.2.5: Loss Curve

This figure the MLPClassifier's loss curve shows how the loss value decreases as the number of training iterations rises. This curve, which has a downward trend suggesting successful training and convergence towards a minimum loss, aids in understanding how well the neural network is learning.

4.3 Discussion

Here's an overall discussion based on the experimental results that show the performance of different classifiers in detecting stress levels across multiple studies.

Classifier Performance: Experimental results show how well different classifiers identify different stress levels. The robustness of decision trees, random forests, and gradient boosting in stress level detection has been shown by their consistently high accuracy in various investigations.

Variability of results: Some classifiers perform consistently across tests, but others have higher variability. Classifier performance is affected by several factors, including feature selection, dataset characteristics, and model hyperparameters.

Effect of model complexity: While advanced classifiers such as gradient boosting and neural network MLP typically produce higher accuracy results, they may require more alignment and processing power. Although simple classifiers such as KNN and logistic regression are easy to understand, their accuracy might drop somewhat.

Evaluation of Metrics: While accuracy gives a general picture of the performance of the classifier, precision, recall, and F1-score provide information on performance in a variety of stress level categories. Taking these indicators into account is essential for evaluating the classifier's accuracy in stress level classification.

Future Research Directions: To further increase the accuracy of stress level identification, further study may examine combination methods, hybrid approaches, or advanced feature selection techniques. Furthermore, examining how environmental, lifestyle, and social variables affect stress level prediction could improve the generalizability of the model.

Real-World Applications: The created stress level detection system has the potential to be used in several fields, such as wellness tracking, healthcare, and stress management treatments. However, before implementation in real-world environments, additional verification and integration with current systems are required.

Future Directions: To increase the accuracy of stress level identification, more studies may concentrate on combination methods or hybrid strategies that combine the advantages of several classifiers. Furthermore, examining feature selection strategies and data beforehand approaches may improve classifier performance and generalizability in a variety of population settings.

All things considered, the experimental design and analysis offer insightful information about the effectiveness of stress level detection classifiers and establish the foundation for further study and useful applications in stress management and well-being.

CHAPTER-5

IMPACT ON SOCIETY, ENVIRONMENT AND SUSTAINABILITY

5.1 Impact on Society

The following are some other effects that stress level-detecting systems have on society:

Health and Well-Being: In today's world, stress is a common problem that can negatively impact both mental and physical health. People can obtain early treatments and assistance by accurately determining their stress levels, which can enhance their overall well-being and health outcomes. This can encourage healthier lifestyles and decrease the stress on healthcare services.

Productivity and Performance: In a variety of contexts, including the workplace and educational institutions, high levels of stress can negatively impact memory, creativity, and productivity. Early stress identification allows employers and educators to put support plans in place for both staff and adolescents, which improves output and performance.

Impact on the Environment: Long-term stress can lead to harmful responses such as excessive resource use, increased waste production, and waste. Stress level detection devices can support environmental conservation efforts and encourage eco-friendly behaviors by addressing stress at its source and encouraging sustainable lifestyle choices.

Social Dynamics: Interpersonal interactions can be impacted by stress, which can result in disagreements, social inequality, and isolation. Stress level monitoring systems have the potential for stronger and more adaptive cultures by promoting empathy, communication, and social connection through enhancing community knowledge and understanding of stressors.

Sustainability: Social and economic well-being are included in sustainable development, in addition to environmental sustainability. Stress level detection systems support the adaptability, resourcefulness, and adaptive capability of individuals, groups, and society by mitigating stress and encouraging holistic approaches to well-being.

Workplace Well-Being: Stress is a common problem that has an impact on workers' health, happiness, and output. Employers can establish a culture of well-being, target interventions, and identify stressors to build better work environments with the use of stress level detection systems. Higher employee satisfaction, retention, and overall corporate performance can result from this.

Education and Learning: Academic pressure, social challenges, and personal issues are just a few of the stressors that students must deal with. Stress level monitoring systems can assist teachers in recognizing children who could be under a lot of stress and in giving them the right kind of assistance and resources. Student progress, enhanced academic performance, and a happy learning environment can all benefit from this.

Community Well-Being: Stress has an effect on people in communities, which affects relationships, social dynamics, and the general well-being of the community. Systems for measuring stress levels can help with community-level programs and Systems for measuring stress levels can help with community-level programs and interventions that target systemic stressors, raise awareness of mental health issues, and develop social support systems. Societies that are healthier and more integrated can result from this by enhancing social bonds and adaptability.

Public Education and Awareness: The use of stress level detection technologies increases public knowledge of the value of stress management and mental health. Through empowering people to make educated decisions about their mental health and fostering a more open and accepting society, these systems educate people about the negative effects of stress on their health and offer useful tools and resources for stress reduction.

Research and Innovation: The creation and application of stress level sensing technologies support environmental psychology and ecotherapy research and innovation. These systems provide insightful information about how stress, mental health, and environmental elements interact. This information can be used to develop eco-friendly behavior advertising initiatives and therapy interventions that use the environment to enhance well-being.

Overall, the implementation of stress level detection systems can indirectly support environmental sustainability by encouraging eco-friendly practices, lowering environmental footprints, fighting for environmental justice, fostering research and innovation, and supporting corporate sustainability initiatives, even though their primary goal is to improve mental health outcomes. These two advantages highlight how closely environmental health and human well-being are related.

5.2 Impact on Environment

Although the primary focus of stress level detection systems is mental health and well-being, there might be substantial environmental consequences associated with their adoption. Here is a more thorough examination of how these systems affect the environment:

Behavioral Shifts: Stress level detection systems frequently work in collaboration with wellness apps that encourage stress-relieving activities like outdoor mindfulness meditation, physical activity, and time spent in nature. Through these activities, people can foster a stronger bond with the natural world and take actions that promote environmental preservation. For instance, people might be more likely to engage in outdoor pursuits like riding, walking, or gardening, which would promote a greater respect for and care for natural areas.

Resource Consumption: Long-term stress can result in unhealthful coping strategies such as excessive waste production, energy use, and product and service consumption. Stress level detection systems decrease the possibility of people participating in environmentally hazardous actions by assisting them in recognizing and managing stressors. These methods contribute to reducing the environmental impact of individuals by encouraging resource conservation and careful consumption.

Environmental Justice Advocacy: People who live in ecologically vulnerable places, such as those who are exposed to pollution, degradation of the environment, or disasters related to climate change, and marginalized communities are particularly impacted by stress. By supporting environmental justice and fair access to mental health resources in impacted communities, stress level detection devices can assist in identifying and addressing the particular pressures experienced by these populations. Through increasing consciousness of the convergence of environmental and social determinants of health, these frameworks support initiatives aimed at tackling systemic disparities and advancing ecological sustainability.

Investigation and Development: The creation and application of stress level sensing technologies support environmental psychology and environmental therapy research and innovation. These systems provide useful data about the connections between stress, mental health, and environmental variables.

Corporate Sustainability Initiatives: Stress leadership and employee well-being programs are becoming a part of many companies' corporate sustainability initiatives. Employers can improve employee satisfaction, productivity, and retention while also supporting more general environmental and social responsibility objectives by putting stress level detection systems into place and providing employees with stress reduction solutions. Workplace wellness programs that promote environmentally friendly transportation choices, eco-friendly office procedures, and green team projects emphasizing community involvement and responsibility for the environment are a few examples of these activities.

All things considered, even if the main goal of stress level detection systems is to enhance mental health outcomes, their application may have a significant impact on environmental sustainability. These systems work to create an improved connection between human well-being and the environment by encouraging eco-friendly habits, fighting for environmental justice, advancing research and innovation, and supporting corporate sustainability initiatives.

5.3 Ethical Aspects

Ensuring the privacy and data security of persons participating in stress level detection activities is a critical function of ethical considerations when working with healthcare facilities, especially Rajshahi Central Hospital. The following are some important moral considerations:

Data protection and privacy: The collaboration includes collecting and examining personal data from people, such as physiological readings, mental health records and behavioral patterns. Strong data security measures such as proper policy and agreements are in place at every stage of the data collection, storage and analysis process to protect people's right to privacy. Users provide their informed consent and the data is carefully managed to avoid misuse or unauthorized access.

Independence and Informed Consent: Participants in stress level detection activities, such as data collection and processing, are given clear and understandable information about the objectives and usage policy, limitations and risks associated with the method. Users are not subjected to excessive argument or pressure in their decision to participate.

Fairness and Bias: Required proper measures are taken to reduce unintentional disadvantages in stress level monitoring systems, such as errors in sample selection, errors in calculation, and errors in cultural assessment tools. This involves verifying assessment tools across different populations and continually evaluating methods to guarantee equal opportunity in stress level assessments.

Stigmatization and Discrimination: Stigmatization and Discrimination: The partnership places a strong emphasis on fostering an environment free from discrimination that values support, sensitivity, and diversity. People are not categorized or stigmatized exclusively because of their mental health issues or stress levels.

Access and Equity: To guarantee equitable access to stress level detection systems and accompanying interventions, consideration is given to variables like economic position, location, and digital literacy. All people, regardless of conditions or background, have access to stress management techniques by addressing the structural obstacles to access and offering alternate ways of access.

Accountability and Responsibility: Accountability and Responsibility: The collaboration maintains moral principles in the creation, application, and use of stress level-detecting technologies. This involves living by the law and professional standards, encouraging an ethical culture, and holding people and organizations responsible for any unethical actions or confidentiality violations.

In summary, ethical standards that give priority to privacy, transparency, fairness, accessibility, and accountability lead to involvement with healthcare facilities like Rajshahi Central Hospital. Ensuring the ethical conduct of stress level detection operations and fostering beneficial results for individuals and society is dependent upon adherence to these essential principles.

5.4 Sustainability Plan

When creating a sustainability strategy for stress detection systems, it's important to take into account how these technologies will affect people in the long run as well as communities and society at large. This is a sustainability strategy designed with stress detection in mind:

User privacy and ethical considerations: It has been made sure that user independence, respect, and privacy are given top priority by stress detection systems. Strong data security procedures are in place to secure private information that was gathered during stress monitoring. Respect towards the laws and ethical standards governing the use of health data, and getting people’s informed consent before gathering or evaluating their stress-related data was ensured.

Assessment of Environmental Impact: How stress detection systems affect the environment over time, taking into consideration the energy usage of data processing facilities and the wearable or monitoring equipment's lifecycle have been examined. Promoting eco-friendly design principles for wearable technology, investigating renewable energy sources for data centers, and optimizing hardware and algorithms to reduce energy consumption.

Fairness and Availability: Inclusivity and fairness were guaranteed through participation of a wide range of people, particularly those from excluded and underserved areas to have fair access to stress detection technologies. Dealing with obstacles to access like money, language, cultural awareness, and computer literacy and working together with politicians, healthcare professionals, and community organizations the advance accessibility and inclusivity in stress detection programs is to be ensured.

Long-Term Health and Well-Being: Making use of stress detection systems is to promote long-term health and well-being results. Place a strong emphasis on developing resilience methods, preventive measures, and comprehensive approaches to mental health that deal with the underlying societal drivers of stress.

Community Empowerment and Engagement: During the development, deployment, and assessment of stress detection systems, involve stakeholders such as individuals, families, medical professionals, and community groups. Encourage a collaborative approach that gives people the confidence to speak up, actively participate in decision-making, and co-create solutions that fulfill their needs and preferences.

Regulatory Compliance and Accountability: Keeping constantly updated on the laws, rules, and moral principles that apply to stress detection technology regulatory compliance is planned to be maintained. These may include rules about data protection, medical device regulations, and professional standards of practice. Make sure that industry standards and legal requirements are followed, and set up accountability systems to handle any ethical or legal transgressions.

Continuous Evaluation & Quality Improvement: Put in place procedures for the continuous assessment and enhancement of stress detection systems' quality. Track the key performance measures for user happiness, accuracy, usability, and health outcomes. To find areas that need work, get input from users and stakeholders. Then, make necessary iterations to the system's functionality, design, and service delivery strategies.

Others can build and implement stress detection systems that support resilience in the community, enhance individual well-being, and ultimately lead to a healthier and more equitable society by incorporating sustainability concepts into the process.

CHAPTER 6

SUMMARY, CONCLUSION, RECOMMENDATION AND IMPLICATION FOR FUTURE WORKS

6.1 Summary of the Study

The study aims to develop and assess stress detection systems through the complete use of many models and attributes. A carefully selected dataset was used, which included many characteristics related to stress levels, such as physiological markers, self-reported symptoms, and lifestyle factors. The research enabled effective data processing and analysis by utilizing Google Colab's computational capacity and collaborative features. The choice and assessment of several machine learning algorithms, such as Support Vector Machines, Random Forest, Gradient Boosting, and Neural Networks, was a crucial component of the research. A strict minimum assurance level of 95% was set for these models' testing and training in order to guarantee the accuracy and dependability of the stress detection results. A variety of performance indicators, such as accuracy, precision, recall, and F1-score, were included in the evaluation process. Graphs, charts, and statistical analysis were used as visualization tools to obtain a more profound understanding of stress patterns and model efficacy. Notably, models with 95% accuracy rates were attained by Decision Trees, Gradient Boosting, and Random Forest. The study also emphasized how crucial it is to keep ethical issues in mind at every stage of the research process. Maintaining user privacy, openness, and equity was of utmost importance, demonstrating a dedication to the conscientious and moral application of stress detection technologies. These guidelines are essential for building user and stakeholder trust and promoting the usage of stress detection systems in practical settings. All things considered, the study's conclusions provide important information about how stress detection systems should be developed and assessed, as well as useful strategies and best practices for obtaining accurate and trustworthy stress assessments. This study establishes a standard for future research in this area by highlighting the importance of multidisciplinary cooperation, moral leadership, and ongoing assessment to progress stress detection technology and enhance general health.

6.2 Conclusion

The study has produced strong proof of the effectiveness of using several machine learning models to reliably identify stress levels based on a wide range of characteristics. Among these, Random Forest, Gradient Boosting, and Decision Trees performed particularly well, obtaining remarkable 95% accuracy rates. These findings highlight the models' potential for use in real-world stress detection scenarios and indicate bright futures for their application in realistic environments. The study's main finding is that stress has several facets, making it necessary to take into consideration a range of variables when developing reliable stress detection systems, including physiological and lifestyle indicators. Through the utilization of an extensive dataset that includes these varied characteristics, the research has illustrated the significance of adopting a complete strategy for stress evaluation, thus improving the precision and consistency of identification results. The study also highlights how crucial it is to take ethics into account while implementing and utilizing stress detection systems. Privacy protection, fairness, and transparency are acknowledged as crucial components in order to guarantee that these technologies are used in an ethical and responsible manner and to foster confidence and trust among users and stakeholders. In summary, this study's findings not only advance our knowledge of stress detection techniques but also offer useful information for the creation of morally and practically sound stress detection systems. To improve health and well-being results for both individuals and communities, more study and cooperation are necessary in the future to refine and optimize these systems.

6.3 Implications for Additional Research

The following research implications are provided in considering of the study's limitations and conclusions:

Investigation of More Complex Models: Subsequent research efforts may go into the incorporation of advanced machine learning models, including collaborative methods, recurrent neural networks, or deep learning architectures, in the context of stress detection. These models could provide better scalability and performance, especially when dealing with large and diverse data sources.

Multimodal Data Incorporation: To improve the precision and depth of stress detection systems, researchers could look at the integration of multimodal data sources, such as social media data, behavioral patterns, physiological signals, and voice analysis. Integrating several data modalities could lead to a more thorough comprehension of each person's unique stress reactions.

Real-time Stress Monitoring: Wearable sensors, smartphone apps, and Internet of Things (IoT) devices may be used in the creation of real-time stress monitoring systems in future research. These technologies might be able to track people's stress levels continuously, which would allow for prompt interventions and individualized support.

Individualized Intervention Techniques: Based on unique stress profiles and preferences, researchers might investigate the creation of individualized intervention techniques. To assist with stress management and developing resilience, this may include the application of adaptive feedback mechanisms, customized recommendations, and behavior modification strategies.

Longitudinal Research and Outcome Measures: To assess the long-term viability and effectiveness of stress detection and intervention strategies, longitudinal research is necessary. Standardized outcome measures may be used in future studies to assess how stress detection systems affect different health outcomes, such as productivity, mental health, and quality of life.

Ethical and Societal Implications: The ethical, legal, and societal consequences of stress detection technology should be further explored in future research. This covers factors including biased algorithms, privacy, equity, and openness. Working together with stakeholders from different backgrounds is crucial to ensuring that these technologies are deployed responsibly and fairly.

Validation and Generalization: To evaluate the durability and generalization of stress detection models across various populations, environments, and situations, more validation research is required. This may include measuring against recognized standards and guidelines, cross-validation across several datasets, and studies of replication in various cultural contexts.

Human-Computer Interaction (HCI) Research: When it comes to creating intuitive user interfaces and interaction models for stress detection systems, HCI research can be quite helpful. Further study attempts may explore human inclinations, usability obstacles, and design ideas to enhance user involvement and adoption of these technologies.

Interdisciplinary Cooperation: To advance the field of stress detection, cooperation between researchers in the fields of psychology, neurology, public health, computer science, and social sciences is crucial. To address complicated issues and create comprehensive strategies for stress evaluation and intervention, future research should promote interdisciplinary collaboration.

Scholars can enhance the understanding of stress detection and management and eventually improve health outcomes and well-being for individuals and communities by addressing these implications in their future research initiatives.

REFERENCES:

[1] A. Hasanbasic, M. Spahic, D. Bosnjic, H. H. Adzic, V. Mesic, and O. Jahic, "Recognition of stress levels among students with wearable sensors," 2019 18th Int. Symp. INFOTEH-JAHORINA, INFOTEH 2019 - Proc., no. March, pp. 20–22, 2019.

[2] S. Betti et al., "Evaluation of an integrated system of wearable physiological sensors for stress monitoring in working environments by using biological markers," IEEE Trans. Biomed. Eng., vol. 65, no. 8, pp. 1748–1758, 2018.

[3] O. Sakri, C. Godin, G. Vila, E. Labyt, S. Charbonnier, and A. Campagne, "A Multi-User Multi-Task Model for Stress Monitoring from Wearable Sensors," 2018 21st Int. Conf. Inf. Fusion, FUSION 2018, pp. 761–766, 2018.

[4] A. Sano and R. W. Picard, "Stress recognition using wearable sensors and mobile phones," Proc. - 2013 Hum. Assoc. Conf. Affect. Comput. Intell. Interact. ACII 2013, pp. 671–676, 2013.

[5] D. Civitello, D. Finn, M. Flood, E. Salievski, M. Schwarz, and Z. Storck, "How do physiological responses such as respiratory frequency, heart rate, and galvanic skin response (GSR) change under emotional stress?" 2014.

[6] A. G. Airij, R. Sudirman, and U. U. Sheikh, "GSM and GPS based real-time remote physiological signals monitoring and stress levels classification," 2nd Int. Conf. BioSignal Anal. Process. Syst. ICBAPS 2018, pp. 130–135, 2018.

[7] B. Alić, D. Sejdinović, L. Gurbeta, and A. Badnjevic, "Classification of stress recognition using Artificial Neural Network," 2016 5th Mediterr. Conf. Embed. Comput. MECO 2016 - Incl. ECyPS 2016, BIOENG.MED 2016, MECO Student Chall. 2016, pp. 297–300, 2016.

[8] Y. Deng, Z. Wu, C. H. Chu, and T. Yang, "Evaluating feature selection for stress identification," Proc. 2012 IEEE 13th Int. Conf. Inf. Reuse Integr. IRI 2012, pp. 584–591, 2012.

[9] M. A. Jiménez-limas, C. A. Ramírez-fuentes, B. Tovar-corona, and L. I. Garay-jiménez, “a physiologycal signals set,” pp. 1–5, 2018.

[10] R. S. Anand and V. Kumar, “Gaurav2017,” pp. 1–5.

[11] A. E. U. Berbano, H. N. V. Pengson, C. G. V. Razon, K. C. G. Tungcul, and S. V. Prado, “Classification of stress into emotional, mental, physical and no stress using electroencephalogram signal analysis,” Proc. 2017 IEEE Int. Conf. Signal Image Process. Appl. ICSIPA 2017, pp. 11–14, 2017.

[12] A. Ghaderi, J. Frounchi, and A. Farnam, “Machine learning-based signal processing using physiological signals for stress detection,” 2015 22nd Iran. Conf. Biomed. Eng. ICBME 2015, no. November, pp. 93–98, 2016.

[13] A. M. Khan and M. Lawo, “Developing a system for recognizing the emotional states using physiological devices,” Proc. - 12th Int. Conf. Intell. Environ. IE 2016, pp. 48–53, 2016.

[14] J. Zhang, W. Wen, F. Huang, and G. Liu, “Recognition of realscene stress in examination with heart rate features,” Proc. - 9th Int. Conf. Intell. Human-Machine Syst. Cybern. IHMSC 2017, vol.1, pp. 26–29, 2017.

[15] L. Vanitha and G. R. Suresh, “Hybrid SVM classification technique to detect mental stress in human beings using ECG signals,” ICACCS 2013 - Proc. 2013 Int. Conf. Adv. Comput. Commun. Syst. Bringing to Table, Futur. Technol. from Around Globe, 2014.

[16] I. C. Valenzuela et al., “Quality Assessment of Lettuce using Artificial Neural Network,” pp. 1–5, 2017.

[17] R. S. Concepcion and L. C. Ilagan, “Application of Hybrid Soft Computing for Classification of Reinforced Concrete Bridge Structural Health Based on Thermal-Vibration Intelligent System Parameters,” Proc. - 2019 IEEE 15th Int. Colloq. Signal Process. its Appl. CSPA 2019, no. March, pp. 207–212, 2019.

[18] M. Šorf, V. Eck, L. Janku, and L. Lhotska, “Application of Neural Network for Stress Classification,” State Art Comput. Intell., pp. 373–375, 2000.

[19] R. G. De Luna et al., “Identification of philippine herbal medicine plant leaf using artificial neural network,” HNICEM 2017 - 9th Int. Conf. Humanoid, Nanotechnology, Inf. Technol. Commun. Control. Environ. Manag., vol. 2018-Janua, pp. 1–8, 2018.

13-7-24

PREDECTION OF STRESS LEVEL BASED ON PHYSICAL ACTIVITY USING MACHINE LEARNING

ORIGINALITY REPORT

12%	12%	2%	9%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	dspace.daffodilvarsity.edu.bd:8080 Internet Source	5%
2	Submitted to Daffodil International University Student Paper	2%
3	Submitted to University of North Carolina, Greensboro Student Paper	1%
4	www.researchgate.net Internet Source	1%
5	www.ijraset.com Internet Source	<1%
6	Submitted to Student Paper	<1%
7	www.itb.ie Internet Source	<1%
8	thailogoprint.com Internet Source	<1%
	Submitted to RMIT University	