

Context–Aware Stress Prediction System

By

Raneem Abdullah Alharthi

Thesis submitted to the
Faculty of Graduate and Postdoctoral Studies
In partial fulfillment of the requirements
For the M.Sc. degree in
Systems Science

School of Electrical Engineering and Computer Science
Faculty of Engineering
University of Ottawa

© Raneem Abdullah Alharthi, Ottawa, Canada, 2016

Abstract

Stress is now recognized as one of the major causes of physical and psychological illness. It is known as a reaction to surrounding environmental threats and the best way to manage it is to understand its triggers. Although people continuously react to their surrounding environments, they sometimes are not aware that certain elements in their environment are considered to be stressors.

Based on this fact, researchers have recently proposed context-aware stress management systems. Most of the proposed systems use context data to provide real time stress monitoring and visualization, along with intervention techniques. However, these interventions are limited to the second and tertiary stages and very little attention has been given to the primary stage.

In this thesis, we introduce a system called CASP. The system's objective is to provide stress status predictions based on a user's current contextual data. Therefore, a detection method is developed using heart rate variability (HRV) as a stress indicator to deliver personalized context-aware stress reports. Based on the predicted status, the system provides users with stress interventions at an early stage in order to help avoid and/or eliminate the occurrence of stress. Our evaluation results show that the CASP system is able to predict the stress status of a user with an averaged accuracy of 78.23% through our limited activity, when compare to a stress status measured using physiological signals. Moreover, it provides prediction models that adapt to the changing nature of both the user's stress status and the surrounding environment.

Acknowledgements

First and foremost, I would like to express my gratitude to God for giving me the health, strength and patience needed throughout this educational journey, allowing me to face all the ups and downs I encountered.

A special thanks goes to my beloved parents, Abdullah and Sabah, for their faith in me and their belief that I could one day have as much success as I do today. They have always been my main source of love, encouragement, support and strength. Everything that I have done in my life so far, I owe to them. A special thanks also goes to my siblings, for always being with me and granting me love, motivation and care. I am also greatly indebted to Rajwa who has constantly provided encouragement, advice and guidance. A heartfelt thanks goes to Raghad for her emotional support; she helped me stay strong through this endeavour. To all of them, I dedicate this thesis.

Likewise, I would like to deeply thank my supervisor, Professor Abdulmotaleb Elsaddik for giving me the opportunity to work in his lab. Thank you for your constant support, advice, encouragement and patience. This thesis would not have been possible without your expert guidance. I would also like to thank Professor Hussein Al Osman, Dr. Mohammed Huda and Dr. Haiwei Dong for sharing their knowledge and providing invaluable feedback.

Contents

1	Introduction	1
1.1	Motivation	2
1.2	Problem Statement	3
1.3	Thesis Contribution	4
1.4	Thesis Organization	4
2	Background	5
2.1	Stress Management Systems	5
2.1.1	Context Awareness in Stress Management Systems	6
2.1.2	Cloud Computing in Stress Management Systems	8
2.1.3	Intervention Methods in Stress Management Systems	8
2.2	Related Works	9
2.2.1	Stress Management Systems for a Specific Context	9
2.2.2	Context-Aware Stress Management Systems	10
2.2.3	Mobile-Based Context-Aware Stress Management Systems	11
2.3	Summary and Comparison	12
3	Context-Aware Stress Prediction System	17
3.1	Proposed System	17
3.1.1	Client Side Components	18
3.1.2	Cloud Side Components	25
3.2	System Flow	31
4	System Implementation	37
4.1	Hardware Components	38
4.1.1	The Biofeedback Sensor	38
4.1.2	Android Device	38
4.2	Software Components	38
4.2.1	Client Side Components	38
4.2.2	Cloud Side Components	46
4.3	CASP System Interfaces	48
4.3.1	Login Interface	48
4.3.2	Personal Information Interface	49

4.3.3	Locating Related Places Interface.....	49
4.3.4	Stress Monitoring Interface.....	49
4.3.5	Stress Intervention Interface.....	50
5	System Evaluation and Experimental Results.....	51
5.1	Evaluation Hypothesis.....	Error! Bookmark not defined.
5.2	Experimental Setup.....	52
5.2.1	Participants.....	52
5.2.2	Experimental Phases.....	53
5.2.3	Apparatus.....	56
5.3	Evaluation of the Experimental Results.....	56
5.3.1	Evaluation Metrics.....	56
6	Conclusion and Future Work.....	68

List of Tables

Table 2.1: A Summary of the Comparison Between the CASP System and the Related Works, Considering Aspects 1 and 2.....	14
Table 2.2: A Summary of the Comparison Between the CASP System and the Related Works, Considering Aspects 3 , 4 and 5.....	15
Table 2.3: A Summary of the Comparison Between the CASP System and the Related Works Considering Aspect 6, 7, and 8.....	16
Table 3.1: Stress Status Classification.....	19
Table 4.1: Linguistic and Numerical Explanation of the Fuzzified Sets of the Distance Between the Current Location of the User and their Home and Work Locations.	40
Table 4.2: Linguistic and Numerical Explanation of the Fuzzified Sets of Weather Conditions.....	41
Table 4.3: Linguistic and Numerical Explanation of the Fuzzified Sets of Days.....	41
Table 4.4: Linguistic and Numerical Explanation of the Fuzzified Sets of Time.....	42
Table 4.5: Users Information.....	44
Table 4.6: Locations Information.....	44
Table 4.7: Users Benchmark Values.....	45
Table 4.8: Standard Deviation and Median Values.....	45
Table 4.9: Context-Aware Stress Status Results.....	46
Table 4.10: The Context-Aware Stress Status Prediction.....	47
Table 5.1: Detailed Description of the Subjects.....	53
Table 5.2: Participant Datasets.....	54
Table 5.3 Attributes Description.....	55
Table 5.4: Prediction Model Accuracy.....	57
Table 5.5:Detailed Accuracy by Class for Individual Prediction Models using Precision, Recall, and F-Measure Parameters.....	60

List of Figures

Figure 3.1: CASP System Components	17
Figure 3.2: Stress Detection Algorithm Flowchart	21
Figure 3.3: The Controller	25
Figure 3.4: The Flow of User Data Through the CASP System.....	31
Figure 3.5: General CASP System Flowchart	33
Figure 3.6: General CASP System Interaction Diagram	34
Figure 3.7: Client Side Interaction Diagram.....	35
Figure 3.8: Cloud Side Interaction Diagram.....	36
Figure 4.1: CASP System Software and Hardware Components	37
Figure 4.2: Zephyr BioHarness.....	38
Figure 4.3: The Entity Relationship Diagram of the CASP System Database	43
Figure 4.4: CASP System Interfaces Architecture	48
Figure 4.5: CASP System Interfaces	49
Figure 4.6: Stress Intervention Interface.....	50
Figure 5.1: Comparison Between Ground Truth Stress Status and Predicted Stress Status for All Users.....	59
Figure 5.2: The prediction Results Classified by the Relaxed Status for All Users	61
Figure 5.3: The Prediction Results Classified by the Stressed Status for All Users.....	61
Figure 5.4: The Prediction Results Classified by the Neutral Status for All Users	62
Figure 5.5: Analysis of the Way Contextual Data Affects the Stress Status of User One.....	63

Figure 5.6: Analysis of the Way Contextual Data Affects the Stress Status for User Two.	64
Figure 5.7: Analysis of the Way Contextual Data Affects the Stress Status for User Three.	65
Figure 5.8: Analysis of the Way Contextual Data Affects the Stress Status for User Four.	66
Figure 5.9: Analysis of the Way Contextual Data Affects the Stress Status for User Five.....	67

List of Acronyms

ANS: Autonomic Nervous System
BITs: Behavioural Intervention Technologies
DT: Decision Tree
ECG: Electrocardiography
EDA: Electrodermal Activity
EMG: Electromyography
FT: Finger Temperature
GSR: Galvanic Skin Response
HF: High Frequency
HMM: Hidden Markov Model
HR: Heart Rate
HRV: Heart Rate Variability
LF: Low Frequency
LR: Logistic Regression
NN: Neural Network
PNS: Parasympathetic Nervous System
PSD: Power Spectral Density
RMSE: Root Mean Square Error
SDNN: Standard Deviation between NN Intervals (Standard Deviation of HRV Signal)
SNS: Sympathetic Nervous System
ST: Skin Temperature
SVM: Support Vector Machine

Chapter 1

1 Introduction

Every day, people face a variety of different challenges and obstacles. The human body has been designed to deal with such challenges and difficult tasks by reacting to threats to the mind or the body. The human being's ability to defend itself and take risks to save its life is due to the Autonomic Nervous System (ANS) [30]. Eustress is a positive kind of stress produced by our bodies to overcome challenges, face our fears, meet deadlines, write exams and speak publicly. Stress can be positive when it helps motivate us and encourages us to take action[47] However, experiencing too much stress for a long period of time can be overwhelming and can have negative effects on our physical and psychological well being. This type of problematic stress is called distress [45].

There are strong associations between stress and certain sever disorders. The American Institute lists more than 50 signs of stress and disorders related to stress[5]. A few examples of such disorders include: headaches, heart attacks, and depression. Moreover, people who experience high levels of stress have greater odds of developing malignant-tumors [14]. Stress can also contribute indirectly to illnesses that are related to unhealthy habits that tend to increase under high stress, for instance smoking and lack of sleep [33].

Whenever a threat is present, human beings can alter their stress status by activating the sympathetic branch of the ANS, which is the branch responsible for triggering the “fight or flight response” [30]. Based on the fact that stress is ubiquitous, threats can be both internal and external. After facing the threat with either a fight or a flight response, the other branch of the ANS (the parasympathetic branch) is activated to calm the body. From the five stressors categories that classified based on duration described in [18] (acute stressors, brief naturalistic, stressful event sequences, chronic stressors, distant stressors), we decided to focus on the acute time limited stressors, as people face this type of stressors more frequently.

The best way to manage stress is to understand its causes. Stress is associated to the relationship between a person and his/her surroundings[30]. The way in which people

have to react and adapt to these surroundings changes whether it is a sudden or a gradual evolution. Some of these changes are considered stressors. As per its definition[30], stress is always related to the body reacting to different kinds of threats, either physical or psychological. Thus, the recent emergence of context-aware stress management systems has been brought forward by the importance of contextual data in stress management.

The existing systems are designed with different stress recognition and management methods, based on their final goal, whether that is stress detection, monitoring and visualizing, or management. They have also been developed with different settings: clinical based or real life. Moreover, with the support of context-aware technology and biofeedback sensors[38], and with the emergence and exponential increase in smartphone ownership, these systems' ability to detect and manage stress has recently improved.

Stress management systems are divided into three main categories based on their stages of intervention [34]. These are: the primary stage, the secondary stage, and the tertiary stage. These three stages contribute differently to the management of people's stress status. The tertiary stage is considered the final stage, where the stress is already harming the body. At this stage, the system helps with the recovery phase. The intervention in the secondary stage deals with current ongoing stress that has been detected and provides techniques to help the individual cope with it. However, the most important and effective stage in which to intervene is the primary stage of stress management, where we can reduce the stress and try to eliminate the stressors before the body is negatively affected [35].

As there is a lack of stress management systems that provide services in the early stage, we propose a context-aware stress prediction system that can provide a stress management services in the primary stage. With the aid of smartphone based context-aware technology, cloud computing and machine learning algorithms, we attempt to predict users stress status based on their current contextual data.

1.1 Motivation

The harmful effects of the daily stress people currently face is due to the limitations of the presently available stress management systems. Acute stress is the type

of stress that can be faced by anyone at any time during the day. The typical clinical based systems are unable to manage the ongoing daily stress people face in their lives. To overcome these limitations, more sophisticated non-clinical based systems are needed. Although many smartphone and context-aware stress management systems are proposed, to the best of our knowledge none of them provide a stress prevention method based on the current context of the user. This motivates us to provide an early intervention method (via the cloud) to enhance the stress management process and eliminate the stressors. One of the principle goals of this work is to show the correlation between a person's surrounding and their stress status in order to illustrate how such data can help predict the stress status. Two main facts are considered as requirements in the design of our system.

1. The changing nature of the environmental context
2. The subjective nature of stress

Hence, the provided prediction models should be adaptive and personalized to meet these requirements.

1.2 Problem Statement

The limitations in the area of stress management systems differ based on the systems goals and requirements. In the following section we present some of these major problems/limitations:

- The first and most important point is the intervention stage of the existing systems. Most of the systems only provide intervention methods in the secondary and tertiary stages. There is a lack of intervention in the early stage, which is also considered the prevention stage[10].
- Contextual information has been used in many context-aware mental healthcare systems to provide suitable services based on the user's current context. However, none of the systems predict the stress status of the user based on the current contextual data.

1.3 Thesis Contribution

The contributions of the thesis are as follows:

- 1 Design and implement a mobile-based Context-Aware Stress Prediction System (CASP):
 - A. Design and implement a context-aware stress status detection algorithm.
 - B. Design and implement a context-based stress status prediction algorithm.
 - C. Design and implement a cloud-based early stress intervention method.

1.4 Thesis Organization

This thesis is organized as follows:

Chapter 2 provides the background and related works.

Chapter 3 presents the context-aware stress prediction system along with its hardware components and software algorithms.

Chapter 4 explains the system's implementation and technical details for hardware and software components.

Chapter 5 presents the evaluation hypothesis, experimental settings and results.

Chapter 6 provides a conclusion and our vision of future work.

Chapter 2

Background and Related Work

This chapter presents the background and related works of stress management systems. The background information is provided in Section 2.1. In Section 2.2, we will discuss the existing systems and highlight their stress monitoring and classification methods. In addition, we will explain how context-aware technology, smartphones, and cloud computing could contribute to the improvement of these systems. Moreover, the role of the intervention methods and their contribution to stress management will also be discussed. Finally, in Section 2.3, a brief comparison of the existing related work systems and of our proposed system will be presented, demonstrating how they differ and how the latter overcomes some of the limitations.

2 Background

2.1 Stress Management Systems

Existing stress management systems all aim to manage stress and reduce its harmful effects on the well-being of humans. The traditional methods of monitoring and managing stress, including interviews, questionnaires, self-reports, and clinic-based measurements, all have to be completed frequently by a physician. These methods are not suitable for real life (non-clinical) use. Such methods can easily miss important information, since the measurements are only taken at limited times and places. There is also a high chance that certain information about the stressors will be neglected, as people might forget to mention elements during the interview sessions [15].

Thus, real life continuous monitoring methods are needed to overcome the limitations of the traditional methods and adapt to real life, which changes rapidly. Stress management systems that use real life measurement methods have the ability to provide long-term stress monitoring and management. These systems are most likely to cover all stressful situations. Therefore, most of the recent stress management systems are using

biofeedback sensors instead of clinical based equipment. There are different types of biofeedback sensors used for the purpose of stress diagnosis, such as a Wristband Skin Conductor, Ambu Blue Gel sensor, a Finger Temperature sensor, and an Electrodermal Activity (EDA) sensor. These sensors are used to detect different physiological stress indicators that are known as bodily reactions to stress. For example, the Heart Rate (HR) and Heart-Rate Variability (HRV) obtained from electrocardiogram (ECG) signals [25] have a strong association with stress levels [9]. Finger Temperature (FT) and Skin Temperature (ST) [42] are other indicators; they show the relationship between decreased body temperatures and stress. The Galvanic Skin Response (GSR) has also been used as an indicator since the skin reacts to stress with increased sweating, which in turn increases the skin's conductivity [44]. Other indicators including breathing rate (Respiration) [42], cortisol level, voice and facial expressions, and sleep quality were also studied and analyzed [13, 23]. These sensors allow users to self-diagnose, and any individual can use them regardless of their background. Users are provided with instructions to help monitor and easily control their stress levels [6].

In order to analyze and classify the stress from the collected indicators, various methods and algorithms are applied such as case based reasoning, fuzzy techniques [1] and machine learning algorithms [11, 39, 49]. In the following sections, we explain how recent technologies could improve real life (non clinical-based) stress management systems.

2.1.1 Context Awareness in Stress Management Systems

The limitations of the clinical based stress recognition methods, as mentioned above, are considered to be the most important problem faced by researchers in the field of stress management. They isolate the user from important factors that might be the main cause of stress or have different direct or indirect relationships with the stressors. The subjective nature of stress and the surrounding environmental factors play an important role in the stress recognition process. Therefore, the surrounding environmental factors should be taken into consideration when determining the stressors of a given user. Context awareness has many definitions, depending on the nature of the sensed objects. In [36], context-awareness has been defined as selected objects examined in a

continuously changing environment. In stress management systems, the selected objects are peoples' stress statuses, which have been studied in different environmental contexts. For example, studies in [7] show the effects of environmental stressors on human behaviour, especially the stress status. In [19], the authors provide a study of stress in different environmental settings. It has also been proven that our physical environment affects our mental health.

Thus, some of the stress management systems use a smartphone's powerful imbedded sensors to detect these contextual data. Researchers in [36] provided three main smartphone-based environmental contexts, where each context is subjected to three changing environments: a computing environment, a user environment, and a physical environment. The last two environments (the user environment and the physical environment) are mostly related to stress management. The user environment describes the real environment surrounding the user as well as his/her situation, for instance the location and the people nearby. The physical environment describes the temperature and lighting and noise levels, all of which can be considered environmental stressors. Works presented in [2, 43] are examples of mobile-based context-aware stress monitoring and management systems.

Real world contextual data are dynamic. To be aware of that continuous change and its gradual effect on stress, the sensed real world features must be expressively represented. One of the most popular methods used to capture the uncertainties of such vague real world data is the fuzzy set theory [32].

The fuzzy set theory has been used in many areas, such as the medical and bioinformatics fields, where incomplete, inconsistent and uncertain data is expected. It provides degrees of membership representation for the context data and uses a concept much closer to that of human thinking than the classical binary representation [31]. It also provides awareness of the continuous environmental changes and of their gradual effect on our well-being. At the same time, continuously obtaining contextual information in an unobtrusive way plays an important role in developing stress management systems. Mäntyjärvi et al. [32] provide an example of a smartphone based stress management system that is based on the fuzzy set theory.

2.1.2 Cloud Computing in Stress Management Systems

Cloud computing is defined as an on-demand service provider, available anywhere and anytime [28]. Cloud-based technology changes and improves health information technology and health care services. Some examples are the Microsoft HealthVault [28] and the Google Health platform [24]. Cloud computing can reduce the requirement for electronic health records, such as personal hardware, software, and networking. Moreover, it has the ability to overcome the computational problems of health related data, using its virtually unlimited computing power. It increases the ability of the remote-monitoring systems by providing a capable storage, where the monitored data are preserved. In addition, it enables real-time healthcare services and interventions by maintaining the records of the patient and preventing their loss.

2.1.3 Intervention Methods in Stress Management Systems

Recently, Behavioural Intervention Technologies (BITs) have received a great deal of attention in mental health. As mentioned in different research on environmental stressors [35], stress intervention methods can help an individual modify his or her behaviour in order to better react in stressful situations. Mobile devices with embedded technologies enhanced these intervention methods. Mobile based intervention methods can be delivered and activated anytime and anywhere using different techniques such as text, video, voice, automated messaging, and web-based and telephone support [6, 12]. These interventions include self-help materials and interactive tools. Of the three different stress intervention methods (primary, secondary and tertiary) classified in [34], today's stress management systems focus on the secondary and tertiary stages. Stress management in the primary stage is still far from optimal because, as presented in [2], there are many requirements for primary stage interventions. In addition to real life monitoring methods, the user must be alerted prior to the occurrence of stress. However, few of the existing stress management systems consider a user's recorded history of mental stress in order to determine the cause of stress [2][36]. Paying less attention to mental stress history data minimizes the system's ability to properly predict the potential occurrence of a stressful situation and identify stressors, which leads to a decrease in the efficiency of the intervention and of the preventive health care techniques.

2.2 Related Works

The existing stress management systems have been developed for a variety of purposes and objectives, using different diagnosis methods and hardware devices. In this section we will present two main categories of stress management systems. First, systems that are designed to cover stress related to specific situations and places. Second, systems that incorporate context-aware technologies. All of the systems will be described based on their final goal, whether that is to provide a stress monitoring service, stress management, or intervention techniques. The mobility and the hardware devices of the system are important aspects that we will also review. Moreover, we will point out the features of each system, such as personalized and adaptive services, that try to meet the subjective nature of the user's stress status. In addition, the intervention methods and their stages will be discussed.

2.2.1 *Stress Management Systems for a Specific Context*

Many stress-monitoring and management systems are designed for specific situations and places. For example:

a) Clinical-Based System

The clinical-based decision support system presented in [1] uses case-based reasoning, textual information retrieval, rule-based reasoning, and fuzzy logic techniques. Finger temperature is used as a stress measurement, and the variation of the temperature is considered a key feature, needed to classify and distinguish between different stress statuses. As an initial implementation, experienced clinicians classified 53 reference cases for the system. The results show the validity of the system compared to that of a human expert.

b) Driving Stress Monitoring System

In [46], the authors present a driving stress monitoring system that aims to reduce the stress related to traffic conditions. They proposed a method based on a correlation analysis with the stress estimation mathematical function. They used ECG, and galvanic skin response GSR physiological signals to extract features for stress level classifications. Their monitoring methodology generates three different stress states (moderate, high, and

low) based on the changing traffic conditions and is able to classify driver stress with an 80% accuracy rate.

c) MoodWings

MoodWings [29] is a wearable biofeedback system designed to remind the driver to calm down and relax during stressful events that occur while driving. Electrodermal activity (EDA) and ECG signal data are gathered using Ambu Blue Sensor gel sensors. The authors designed a butterfly wing motion as a physical interface that is activated to alert the driver of their stress level.

2.2.2 Context-Aware Stress Management Systems

For real life and long-term stress management, being aware of surrounding stressors is considered to be the main key to successfully managing stress. For this reason, some systems recognize and incorporate different users' contexts.

d) A Context-Aware System for Life Events

Personalized contexts of life events were detected with the arousal information of 21 users [26]. Users wore a wristband in order to detect and monitor their physiological signals, based on the skin's conductance features. The collected data were analyzed to find the correlation between the context and the user's reaction, which was to be displayed using a designed software tool. The goal of this work was to draw attention to the user's arousal state as it was affected by daily life events, determined by the user's calendar information.

e) Activity-Aware Stress Monitoring System

Sun et al. [49] present an activity-aware mental stress monitoring system designed based on the fact that certain physical activities might affect mental stress status measurements. ECG, GSR, and accelerometer data were collected in different stages of physical activities. The experiment was conducted for three activities: sitting, standing, and walking. Their classification methods were developed using a combination of machine learning algorithms including machine learning J48 Decision Tree, Bayes Net,

and Support Vector Machine (SVM). This system achieved an accuracy of 92.4% for mental stress status classification.

2.2.3 Mobile-Based Context-Aware Stress Management Systems

With the maturation of mobile and smartphone technologies, many stress management systems now utilize smartphones to take advantage of features like mobility and a variety of embedded sensors. For example:

f) deStress

The deStress system presented in [52] is a mobile-based health care system for remote stress monitoring and management. The authors designed stress monitoring and alleviation algorithms in addition to a biofeedback method. Two different sensors were used for stress detection: a tri-axial accelerometer and the pulse wave sensor pulsometer. The system continuously assesses the user's stress level by detecting speaking and movements. It uses a cloud-based service to share the collected and analyzed data with medical specialists.

g) Mobilyze

Burns et al. present a mobile-based application called Mobilyze, which predicts cognitive states and mental health-related stress states, such as mood and emotions, based on the user's self-reports and data collected using the smartphone's embedded sensors. The system delivers a momentary intervention via a website, where the user can learn more about their behaviour and be connected with their clinician via email/phone calls, if further help is needed. Machine learning models were used for learning purposes. As a result, the application achieved 60% to 91% accuracy when predicting the context data, despite the mood predictions, which generally cause lower accuracy rates.

h) Mobile Heart Health

Similarly, in [43] Intel presented a context-aware prototype called Mobile Heart Health, which can detect and manage stress. The system analyzes the root mean square of successive differences RMSSD of HRV using ECG signals and a mobile-based mood map for regular self-reporting. If the user is stressed, the system provides psychology-based feedback to reduce stress, for instance breathing exercise techniques.

i) U-Biofeedback Reference Model

The authors designed a game using the U-Biofeedback reference model on the Android platform. The application uses the HRV signal via a Zephyr BioHarness sensor to detect stress levels, and mathematical methods were developed to monitor the stress intensity level. The stress state and its intensity are displayed to the user in real time using a tree object. If the user is relaxed, the tree will grow and produce flowers. However, it will deteriorate if the user is experiencing a stressful event. Also, it has a score that reflects the health improvements of the tree over time. The application provides the user with context-aware and personalized relaxation techniques based on the location context data, user preferences, and the efficiency of the previous techniques. The application reached an accuracy of 89.63% for the correct identification of acute stress.

j) Context-Aware Chronic Stress Recognition System

In [39], the authors present a chronic stress measurement method for everyday situations. It is a context-aware stress monitoring system using embedded mobile sensors. They also used a questionnaire in order to evaluate the user's ability to estimate their stress, as the system aims to provide the users with an awareness of their stressful situations. The system recognizes the user's location, activity, and ambient surroundings using different mobile sensors and matches these data with the user's self-reports. They train and test Hidden Markov Models (HMMs) for the classification. The system allows users to visualize their monitored stress, which leads them to learn how to deal with or avoid these situations in the future, in order to reduce the health risks.

2.3 Summary and Comparison

A comparison of the related work systems is shown in Table (2.1) ,Table (2.2) and Table (2.3) based on the following aspects, which describe the main features of the previous stress management systems, and how the CASP system differs from all of these systems:

1. Main goal of the system:
 - Stress monitoring
 - Stress management
 - Stress prediction

2. Stress detection methods:
 - The hardware sensor used to detect and recognize stress and the targeted stress indicators.
 - Stress self-reports
3. Context awareness:
 - Does the system incorporate context-aware technology in general?
 - Is the system designed to manage stress related to limited locations or activities?
4. Mobility:
 - Does the system provide a smartphone-based service?
5. Classification methods
6. Does the system provide a stress intervention method? If so, at what stage is the method provided?
 - Early stage
 - Intermediate stage
 - Late stage
7. System features:
 - Personalized: does the system provide a personalized service?
 - Adaptive: does the system have the ability to adapt to the rapidly changing nature of stress and stressors?
8. System result

Detecting contextual data has an enormous effect on the result of stress management in a mobile health-monitoring scenario. The CASP system differs from all of the presented studies in the way that it enhances the stress prevention method. In our proposed system, we give more attention to the relationship between the contextual data and the user's stress status. We hypothesize that the contextual data provide the necessary auxiliary information to indicate and specify personalized stressors. This contextual data can be used further for the prediction, intervention and avoidance of the stressful events.

Table 2.1: A Summary of the Comparison Between the CASP System and the Related Works, Considering Aspects 1 and 2.

Reference Work	Main Goal			Detection Method	
	Monitoring	Management	Prediction	Hardware Sensor	Self Reports
(a)	✗	✓	✗	Finger temperature	✗
(b)	✓	✗	✗	Galvanic skin response	✗
(c)	✗	✓	✗	Ambu Blue Sensor gel sensors	✗
(d)	✓	✗	✗	Wristband and skin conductance features	✗
(e)	✓	✗	✗	ECG, GSR, and accelerometer	✗
(f)	✓	✓	✗	Tri-axial accelerometer and the pulse wave sensor pulsometer.	✗
(g)	✓	✓	✓ (Using-self reports)	38 smartphone embedded sensors including: global positioning system, ambient light and recent calls.	✓
(h)	✓	✓	✗	Wireless ECG sensor, a Tri-axial accelerometer and RMSSD	✓
(i)	✓	✓	✗	Zephyr BioHarness	✗
(j)	✓	✗	✗	Presence service: Open Mobile Alliance (OMA)	✓
CASP System	✗	✗	✓	Zephyr BioHarness and smartphone sensors: GPS sensor, accelerometer	✗

Table 2.2: A Summary of the Comparison Between the CASP System and the Related Works,
Considering Aspects 3 , 4 and 5.

Reference Work	Context-Aware		Mobile-Based	Methods
	General Context	Limited Context		
(a)	✗	✓	✗	Case-based reasoning
(b)	✗	✓	✗	A correlation analysis method
(c)	✗	✓	✗	Alert method using physical interface
(d)	✓	✗	✗	Correlation method
(e)	✗	✓	✗	Combination of machine learning algorithms
(f)	✗	✓	✓	Stress Monitoring Algorithm Alleviation Algorithm
(g)	✓	✗	✓	Machine learning models
(h)	✓	✗	✓	RMSSD, A mobile-based mood map
(i)	✓	✗	✓	Mathematical methods
(j)	✓	✗	✓	Hidden Markov Models
CASP System	✗	✓	✓	Naive Bayes classifier, Fuzzy Logic for context classification and cloud based prediction models.

Table 2.3: A Summary of the Comparison Between the CASP System and the Related Works
 Considering Aspect 6, 7, and 8.

Reference Work	Intervention			Features		Results
	Early "Prevention"	Intermediate	Late	Personalized	Adaptive	
(a)	✗	✗	✓	✗	✓	81% correctly classified cases compared to an experienced clinician.
(b)	Not Applicable				✓	The system achieved 80% accuracy.
(c)	✗	✓	✗	✓	✓	Positive results from qualitative feedback and satisfaction survey. Mean of the users' agreement = 4.7
(d)	Not Applicable			✓	✓	Based on qualitative evaluation: participants (12 out of 15) increased their self-awareness of stress patterns.
(e)	Not Applicable			✓	✓	The system achieved 92.4% accuracy.
(f)	✗	✓	✗	✓	✓	Based on participants' evaluation, the stress level drops noticeably after alleviation techniques.
(g)	✗	✓	✗	✓	✓	The system achieved 60% to 91% accuracy.
(h)	✗	✗	✓	✓	✓	The mobile therapies successfully displayed the breathing exercise in response to the detected stress.
(i)	✗	✓	✗	✓	✓	89.63% correct identification.
(j)	✗	✗	✓	✓	✓	The users could recognize their ongoing stressors successfully.
CASP System	✓	✗	✗	✓	✓	The system achieves an average of 78.32% of the user's stress status.

Chapter 3

3 Context-Aware Stress Prediction System

In this chapter, we describe the proposed system, which is designed based on the requirements provided in section 1.1. In Section 3.1 the system components is discussed with the explanation of the hardware components and the designed software algorithms. Section 3.2 presents the system flow, the main flowchart of the CASP system, as well as the interaction diagrams.

3.1 Proposed System

The proposed CASP system consists of two main components, namely the client side and the cloud side, shown in Figure (3.1).

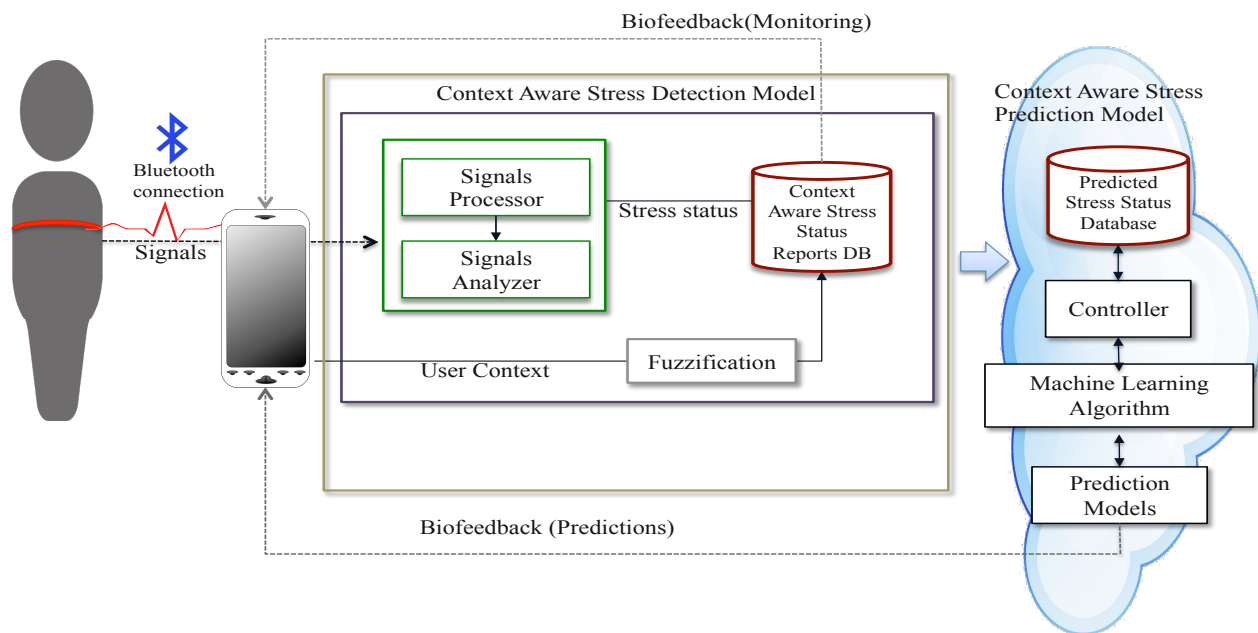


Figure 3.1: CASP System Components

The client side is mainly used for real-life context-aware stress status data acquisition from users, while the cloud side is responsible for generating adaptive and

personalized prediction models. Each component will be described in the following sections.

3.1.1 Client Side Components

3.1.1.1 Biofeedback Sensor

A wearable sensor Zephyr Bioharness [51] is used to extract and send relevant physiological information to the targeted computing device via Bluetooth technology, where the measured signals can be analyzed and processed.

3.1.1.2 Mobile Device

As we aim to provide an "everywhere/anytime" context-aware stress prediction service, we take advantage of smartphones that are capable of providing the functionalities required to detect the contextual data. The recent advances and growing popularity of this device is due to the fact that it has become an important source of information and services. Moreover, it has the ability to connect to different devices and sensors via Bluetooth technology, and use different API's provided by android SDK. By using a variety of imbedded sensors such as GPS, accumulators, and magnetometer, we can detect the user's current time, day, location, and weather conditions.

3.1.1.3 Context-Aware Stress Detection Model

We have designed and developed the Context-Aware Stress Detection Algorithm to calculate the HRV by monitoring the ECG signal. This method is known as the best way to detect stress levels [20] and HRV is considered the standard measurement and physiological interpretation for clinical use. Stress influences the sympathetic nervous system, and the LF/HF ratio parameter changes accordingly. It decreases when a person is relaxed, remains stable in a non-stressed situation, and increases during a state of stress. We derived the three equations in Table 3.1 and Algorithm 1 based on the known effects of the autonomous nervous system on LF/HF, in order to detect the stress status as per this research papers showing the effective usage of the LF/HF signal[8, 16, 22, 30, 41].

To detect the stress levels, we first need to collect benchmarks of the LF/HF ratio. This process takes five minutes, based on the accepted short-term duration of the HRV measurements. During this time, the user's stress level should be moderate, which means somewhere between the no stress state and the stress state. We assume that the

measurements will follow the normal distribution with the following median and standard deviation respectively: $\mu_{LF/HF}$, $\sigma_{LF/HF}$. The Normal Values Range will be calculated for this parameter in order to detect stressful events. The benchmarks will then be used to compare against the LF/HF value, which is generated every 30 seconds, to classify the stress levels into one of the three different levels: stressed, relaxed, or neutral, the latter being defined as indicating to the state where the user's stress status is not stressed and not relaxed, as shown in Table (3.1). The contextual data detected using the smartphone will be updated with any changes and saved every 30 seconds, along with each measured stress status. The algorithm flowchart is shown in Figure (3.2).

Table 3.1: Stress Status Classification

Stress Status	Equation used for classification
Relaxed	$LF/HF (i) < \mu_{LF/HF^*} - \alpha \sigma_{LF/HF^*}$
Neutral	$\mu_{LF/HF^*} - \alpha \sigma_{LF/HF^*} \leq LF/HF (i) \leq \mu_{LF/HF^*} + \alpha \sigma_{LF/HF^*}$
Stressed	$LF/HF (i) > \mu_{LF/HF^*} + \alpha \sigma_{LF/HF^*}$

Where, α is the sensitivity of measuring the stress status changes, and α is =1, which means it is very sensitive to any change in the stress level.

LF/HF (i) is the value of the LF/HF parameter.

LF/HF* is a five minutes duration of the benchmark value collected from the LF/HF parameter as a one time process.

Algorithm 1: Context-Aware Stress Detection

Input

N → User's name

A → User's age

G → User's gender

LF/HF (*i*) → The value of the LF/HF parameter, which is detected every 30 seconds.

LF/HF* → The value of the benchmarks from the LF/HF parameter

K → User's context {C1: Time of day, C2: Day of week, C3: Activities, C4: Weather, C5: Location, C6: Distance from home, C7: Distance from work}

Output

Ss → Context-aware stress status

Procedure

1. *Get* N
2. *Get* A
3. *Get* G
4. *Get* LF/HF*
5. *Calculate* the standard deviation for LF/HF* signals → σ_{LF/HF^*}
6. *Calculate* the median for LF/HF* signals → μ_{LF/HF^*}
7. *Get* K
8. *Get* LF/HF (*i*)
9. *If* $\mu_{LF/HF^*} - \alpha\sigma_{LF/HF^*} \leq LF/HF (i) \leq \mu_{LF/HF^*} + \alpha\sigma_{LF/HF^*}$
10. *Then* Ss → Neutral
11. *Else if* $LF/HF (i) > \mu_{LF/HF^*} + \alpha\sigma_{LF/HF^*}$
12. *Then* Ss → Stressed
13. *Else* $LF/HF (i) < \mu_{LF/HF^*} - \alpha\sigma_{LF/HF^*}$
14. *Then* Ss → Relaxed
15. *End if*
16. *End*

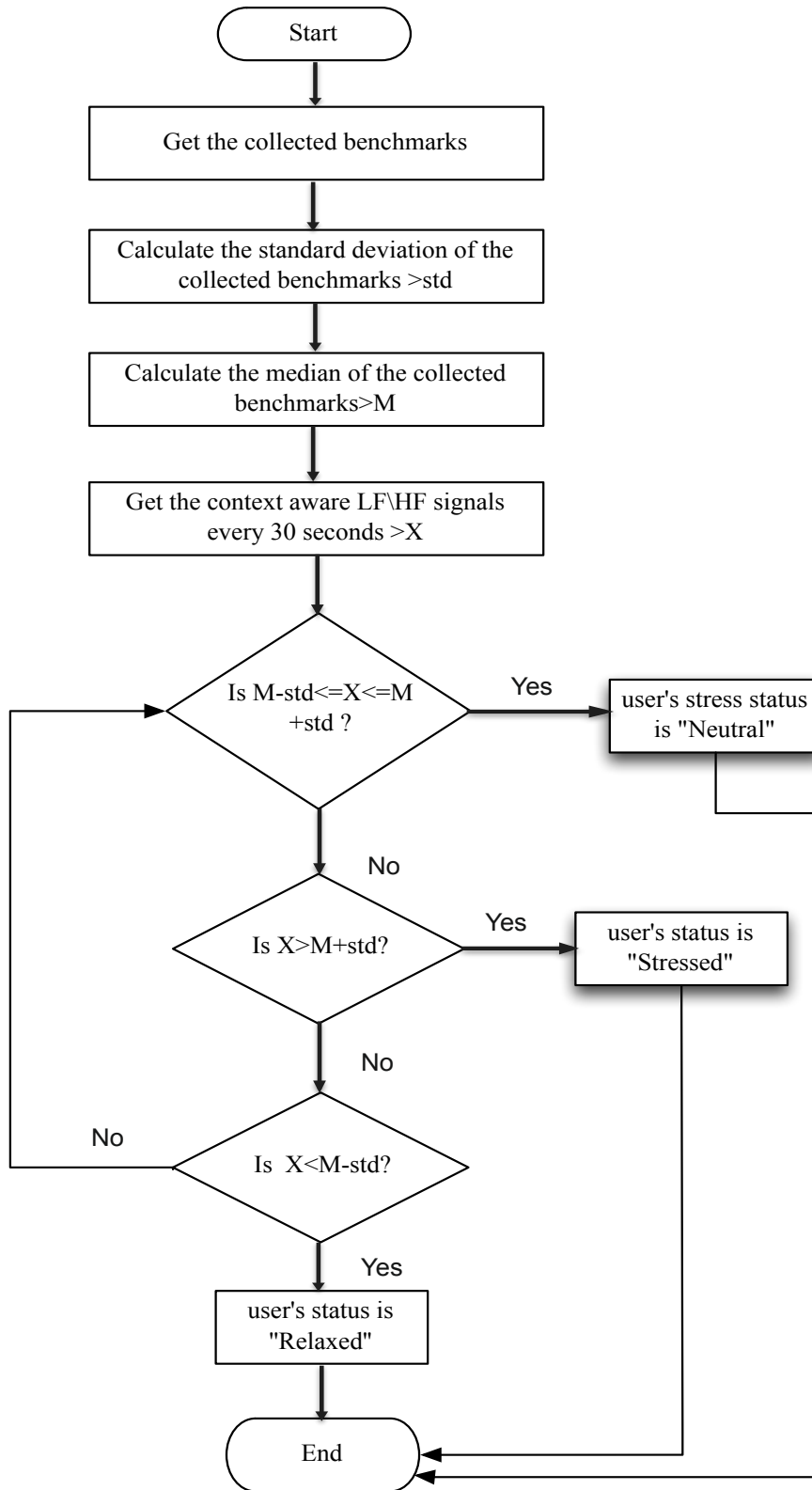


Figure 3.2: Stress Detection Algorithm Flowchart

3.1.1.4 Signal Processor

The signal processor is responsible for performing a variety of signal processing algorithms on the collected sensory data such as: amplification, isolation and multiplexing, as described in [2]. It also does the signal filtering to remove the unwanted components from the measured signals and then uses the linearization methods to correct the nonlinear response of the biofeedback sensor. Most importantly, it converts the analyzed data from the time domain to the frequency domain.

3.1.1.5 Signal Analyzer

The signal analyzer is used to analyze the collected processed signals to obtain the desired output. For example, during the monitoring of the HRV signals, the LF\HF is analyzed by the signal analyzer to classify the user's stress level. Because the signal changes over time, a time domain analysis is required. Nonetheless, the most important parameters of the ECG are more effective in the frequency domain, where it renders more valuable information, and that will be done using the signal processor.

3.1.1.6 Contextual Data Detection

The environmental context might include independent stressors, or contain certain stressors that can lead indirectly to a stressful situation. We have used a variety of embedded sensors in the mobile device along with different API's to collect the following contextual data. The majority of the contextual data is collected automatically without the user's attention, with the exception of the current activity, which requires the users to enter their activity.

- 1 User's current location: this is the location of the user, obtained from their mobile device. We aim to calculate the distance between the current location of the user and the most relevant locations. These locations are selected based on their high likelihood of affecting the user's stress status. To calculate the distance between the user's current location and (home, work) locations, we use the Haversine formula (1):

$$\begin{aligned} a &= \sin^2\left(\frac{\Delta\varphi}{2}\right) + \cos\varphi_1 \cdot \cos\varphi_2 \cdot \sin^2(\Delta\lambda/2) \\ c &= 2 \operatorname{atan} 2(\sqrt{a}, \sqrt{1-a}) \\ D &= R \cdot c \end{aligned} \tag{1}$$

Where:
 $\lambda = \textit{longitude}$
 $\varphi = \textit{latitude}$
R= earth's radius

- 2 Current time of the day is another important aspect of the contextual information, as stressful events (or any events in general) occur repeatedly at certain times.
- 3 Identifying the day of the week helps recognize the day that users perform stressful activities.
- 4 Current weather in the city will track the weather conditions of the user's locations, which might be a potential stressor.
- 5 The current activity of the user shows the kind of activity that the user is performing. We focus on the most common activities such as eating, relaxing, driving, attending a meeting and working, which are provided in the form of a list from which the user can select. The kind of activities that are considered to be high-level activities are those that are difficult to detect with sensors.

A. Contextual Data Fuzzification

Because real world data are not absolutely crisp data, we are converting the detected contextual data from discrete to continuous form using the fuzzy set theory [40]. This theory has been used in different domains to simplify the way uncertain and vague data is handled. It also handles partial truth data that is dependent on human and common sense reasoning to define vague medical entities and their related contextual data as fuzzy sets. We use the technique of fuzzy membership functions, which is defined as degrees of membership representation of data, unlike the classical binary representation. This enables us to be aware of the continuous changes in the dynamic real world and the surrounding environment and how they affect our stress levels, which are gradually affected by the sensible factors.

B. Comparison between the classical and the fuzzy sets

To compare between the classical and the fuzzy sets, we consider time of the day as an example. The crisp value of time "23:59 p.m." will be "night" with a probability equal to one. On the other hand, in fuzzy logic we are dealing with degrees instead of crisp values.

Considering the same example, where the fuzzy value of time is “23:59 p.m.” will be "night" with a degree equal to 0.1 and "midnight" with a degree of 0.9. If we are using crisp values, we would end up having discrete numbers describing each context with sharp boundaries, but with fuzzified context we have linguistic variables with different degrees of membership functions that belong to more than one set and have smooth boundaries. Therefore, some elements can be considered as partial members of two different sets respectively. For instance, the weather data can be 0.8 sunny and 0.2 cloudy.

In a case where the user’s stress level is more affected by the cloudy weather condition, the stress level will increase, as it gets cloudier.

C. Context classification steps

1. All of the input data are converted to the associated linguistic variables.
2. Select one of the membership functions. We have selected the triangular membership function (2) because of its significant and dynamic variation results, which are processed quickly.
3. Calculate the output using one of the fuzzification operations. We are using a union operator (3) to compare the results.

These steps have been completed using the following equations:

- Triangular membership function (2) is used to calculate the membership function

$$\mu(x; a, b, c) = \begin{cases} 0, & x \leq a \\ \frac{x-a}{b-a}, & a \leq x \leq b \\ \frac{c-x}{c-b}, & b \leq x \leq c \\ 0, & c \leq x \end{cases} \quad (2)$$

Where $x =$ *The current value*, and a, b, c are the coordinates of the three corners of each triangular membership function.

- The union set operation (3) is used to get the largest membership value of the elements in either set:

$$\mu A \cup B(x) = \max(\mu A(x), \mu B(x)) \quad (3)$$

3.1.1.7 Context-Aware Stress Status Report Database

The context-aware stress status reports of the users will be saved in an organized way in the local database with the personal information of each user. The records are managed for further use in the prediction model. A new stress status report will be saved every 30 seconds, along with the contextual data that will be continuously updated in order to take into account any changes.

3.1.2 Cloud Side Components

3.1.2.1 The Controller

The controller is responsible for preparing the context-aware stress status reports to be used by the machine learning models. This is done by dividing the personalized reports to different datasets, which will be used in the prediction process. The controller receives the prediction requests and separates the context-aware stress status reports into new and teaching datasets, as shown in Figure (3.3). The teaching dataset is the historical data that was retrieved earlier and is used to build and re-build the prediction models for each user, while the new dataset is the data that came from the user in the prediction request form. For example, a labeled instance will be sent as a prediction request to the controller in the cloud side. The controller will save this instance in the teach dataset, to be used in the next prediction request to re-build the prediction model. For the current prediction, only the attributes will be saved in the new dataset to predict the user's stress status. The result of the prediction will be compared with the actual stress status that was saved earlier in the teaching dataset.

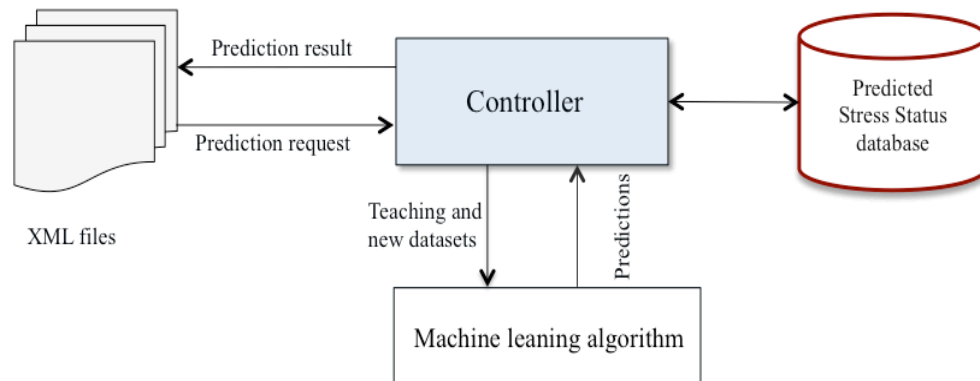


Figure 3.3: The Controller

We have divided the datasets into teaching and new datasets. The teaching dataset includes the labelled user's stress status with the contextual information, while the new dataset only includes the contextual data.

3.1.2.2 Predicted Stress Status Database

We use a remote database that provides functions for secure data storage. This database is used to save all of the predictions that have been sent to the user so that they can be retrieved and used in the future as a teaching datasets by the machine-learning system in the cloud.

3.1.2.3 Machine Learning System

We use a machine-learning algorithm to describe the relation between the user's stress status and the contextual data that represents the possible stressors. Machine learning has been largely applied to analyze data in various domains, but it is still new to personalized mental healthcare services [48], [21].

A. The Context-Aware Stress Dataset

The use of contextual data as features for the prediction model is based on the fact that environmental and surrounding contexts can trigger a user's fight or flight response. Thus, by using these context data we will be able to identify the events or situational factors causing these reactions. Each sample in the stress dataset is described by five features, which are the contextual data described in Section 3.1.1.6.

B. Naive Bayes Classifier

The Naive Bayes algorithm (NB) has been selected as the stress predictor for our system. This algorithm is known as the most effective classification algorithm for medical datasets. An experimental comparison between NB and five other popular algorithms Logistic Regression (LR), KStar (K*), Decision Tree (DT), Neural Network (NN) and a simple rule-based algorithm (ZeroR), have been done on more than 15 medical datasets such as: diabetes disease, breast and lung cancer, biomedicine, heart disease, and lymphography. Most of these datasets have both nominal and numeric attributes [3]. The comparison shows the high performance of NB compared to the other classifiers. The data yielded by this comparison provided us with a strong reason to use the selected classifier.

This classifier has the ability to deal with the characteristics of medical data. The following features of NB are highlighted based on medical dataset requirements, which are close enough to our dataset requirements. First, real-life medical experiments have a high chance of missing values, which is considered a main problem faced by the classifiers, and has been solved by the NB classifier. Second, it has the ability to deal with small amounts of data, which we tend to have in our system at the very beginning, although this data will increase over time. Third, the natural interpretation and explanation ability provided by this classifier is required in the medical field where reasons and causes are considered important information. Fourth, and most importantly, it can provide an explicit explanation of the correlations between the user's stress status and the environmental context [27].

The NB classifier works based on the Bayesian theorem shown in Equation (4):

$$P(H|X) = \frac{P(X|H)P(H)}{P(X)} \quad (4)$$

Here, $P(H|X)$ is the posterior probability of H conditioned on X , and the probability $P(H)$ is called the prior probability of H . The algorithm learns from the training dataset, which contains records of the attributes assigned to the class labels. The tuples, known as records, are assigned to classes by attribute. For example, a tuple X represents the n attributes vector $X = (x_1, x_2, \dots, x_n)$. Possible classes for tuple X are represented by C_1 to C_m . For the purpose of categorization, each tuple X is assigned to one of the C classes. The probability of X belonging to class C is calculated by the algorithm. For labeling, more than one class will be ranked, and the highest labelled class will be selected.

The formal Naive Bayesian classification method is presented in (5):

$$P(X|C_i) = \prod_{k=1}^n P(x_k|C_i) = P(x_1|C_i) \times P(x_2|C_i) \times \dots \times P(x_n|C_i) \quad (5)$$

The classifier used in algorithm (2) assumes the independence of the attribute values of the same class. For example, even though multiple stressors could affect the user's stress status together as one event, as they depend on each other to describe stressful situation factors, they are still used independently for the classification. The attributes are also used as independent factors when calculating the probability. This method is known for its fast

performance on large and small datasets as well as for its simple implementation, which produces a high level of accuracy due to the assumption of class independence.

Ss_j → Represents the stress status

C → Represents the contextual data

$P(Ss_j|C)$ → Represents the probability of being under j stress status in the context C

The class Ss : $P(Ss_j|c)$

User's current stress status= $argmax_j$ in (6)

$$argmax_j = 1 \dots 5p(Ss_j|C_1 \dots C_n) \quad (6)$$

Thus, user's current stress status= $argmaxP$ in (7)

$$argmaxP(Ss|C_1)P(Ss|C_2)P(Ss|C_3)P(Ss|C_4)P(Ss|C_5) \quad (7)$$

All the attributes are conditionally independent.

Given a training dataset D of N labeled examples:

1. Estimating the $P(Ss_j)$ for each class Ss_j is calculated using (8)

$$\hat{P}(Ss_j) = \frac{N_j}{N} \quad (8)$$

N_j → Number of the examples of the class Ss_j

2. Estimate $P(C_i = c_k|Ss_j)$ for each value c_k of the attribute C_i and for each class Ss_j

For the discrete values of the C_i attributes Equation (9) is used.

$$\hat{P}(c_i = c_k|Ss_j) = \frac{N_{ijk}}{N_j} \quad (9)$$

N_{ijk} Number of examples of the class Ss_j having the value C_k for the attribute C_i

And for the continuous values of the C_i attributes

We use the normal distribution in (10)

$$P(C_i = c_k|Ss_j) = g(c_k; \mu_{ij}, \sigma_{ij})g(c, \mu, \sigma) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(c-\mu)^2}{2\sigma^2}} \quad (10)$$

The mean μ_{ij} and the standard deviation σ_{ij} are estimated from D .

Algorithm 2: Generating Adaptive Prediction Models

Input

Pr → Prediction requests

User ID

Output

Ss* → Predicted stress status

Procedure

1. *Get* Pr
3. *Get* user ID
4. *Retrieve* the user's data from database to the controller → Teaching dataset
6. *If* teaching data is enough to build the prediction model
7. *Then*
8. *Do* build prediction model for the user
9. *Make* prediction
10. *Send* back predicted stress state
11. *Save* result in database
12. *Else*
13. *Send* an error message and require more teaching data (instances)
14. *End*

3.1.2.4 Adaptive Stress Prediction Model

We obtained a Naive Bayes prediction model that will be able to provide predictions about the user's stress status based on the environmental context. This model shows what factors are responsible for the user's stress status by marking those factors as stressors. The prediction model is personalized for each user, as stress measurements vary from one user to another. Also, it is an adaptive model that will rebuild each time new data is saved to the database. The new predictions that are delivered to the user are saved to the database to be used as teaching data to build new models. This is based on the fact that the future stress status is more likely affected by the current state [4].

Algorithm 3: The Stress Prediction Algorithm

Input

$S_s \rightarrow$ Measured stress status

$K \rightarrow$ User's context $\{C_1, C_2, C_3, C_4, C_5, C_6, C_7\}$

Output

$S_s^* \rightarrow$ predicted stress status

Procedure

1. *If* the user is not already registered
2. *Then*
3. *Do* start the registration process in algorithm 1
4. *Else*
5. Sign in
6. *Get* K, S_s
7. *Send* prediction request to the cloud
8. *Receive* prediction result
9. *If* the prediction model requires more instances
10. *Then present* the monitored user's stress status
11. *If* the monitored stress status = stressed
12. *Then* start stress management recommendation process
13. *Else if* the predicted stress status = stressed
14. *Then* go to 12
15. *Else*
16. *Send* the stress status
17. *End*

3.2 System Flow

As an overview, the context-aware stress prediction (CASP) system’s architecture is described below:

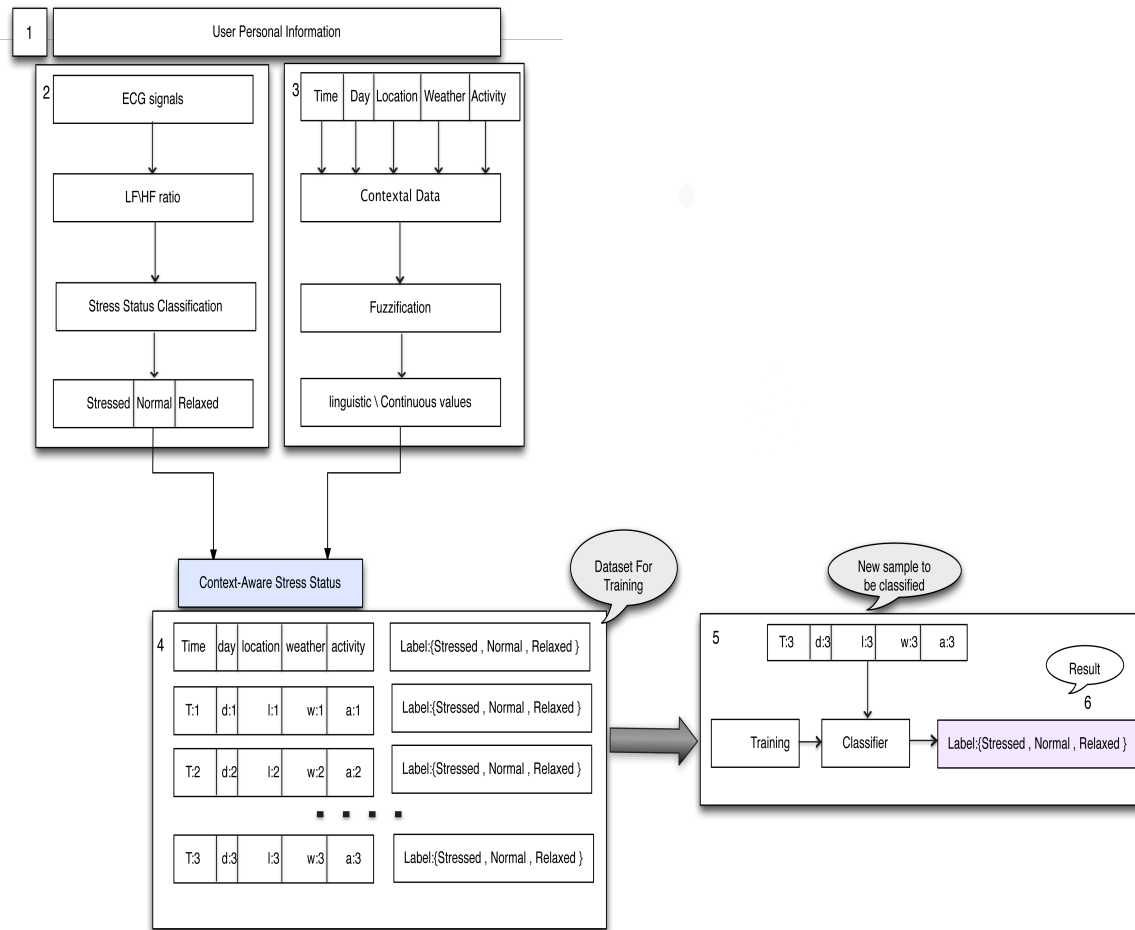


Figure 3.4: The Flow of User Data Through the CASP System

1. As shown in Figure (3.4), the system starts with a registration process to save the user's personal information (full name, age, gender).
2. The user wears a biofeedback sensor, an electrocardiogram (ECG) to measure their stress status. The user's smartphone device must be connected to the Internet and to the biofeedback sensor via Bluetooth, and its location services must be enabled. We measure the user's Heart Rate Variability (HRV) signals, which are among the most popular and commonly used measures of the autonomic nervous system [37]. The HRV signals allow

us to recognize the low frequency/high frequency (LF/HF) power ratio. This physiological parameter has been used clinically to determine stress related disorders. This parameter changes according to the user's stress status; in general, it increases with stress.

A stress status (relaxed, neutral, or stressed) is inferred from the sensor.

3. The detected stress status will be saved along with the context-aware data detected from the mobile device. This includes the user's current location, activity, day, time and weather, as well as the distance from the current location to the user's home and work. The detected contextual data are converted from discrete to continuous values using fuzzification techniques.

4. At 30-second intervals, stress measurements are formulated by the 6 contextual data as features that will be saved in a stress state records database.

5. Prediction requests are then sent to the cloud server, where the personalized and adaptive prediction models are generated using machine learning algorithms.

6. Finally, based on the predicted stress status, the system will send a relaxation technique to the user as an early stress intervention. The general system flowchart is shown in Figure (3.5).

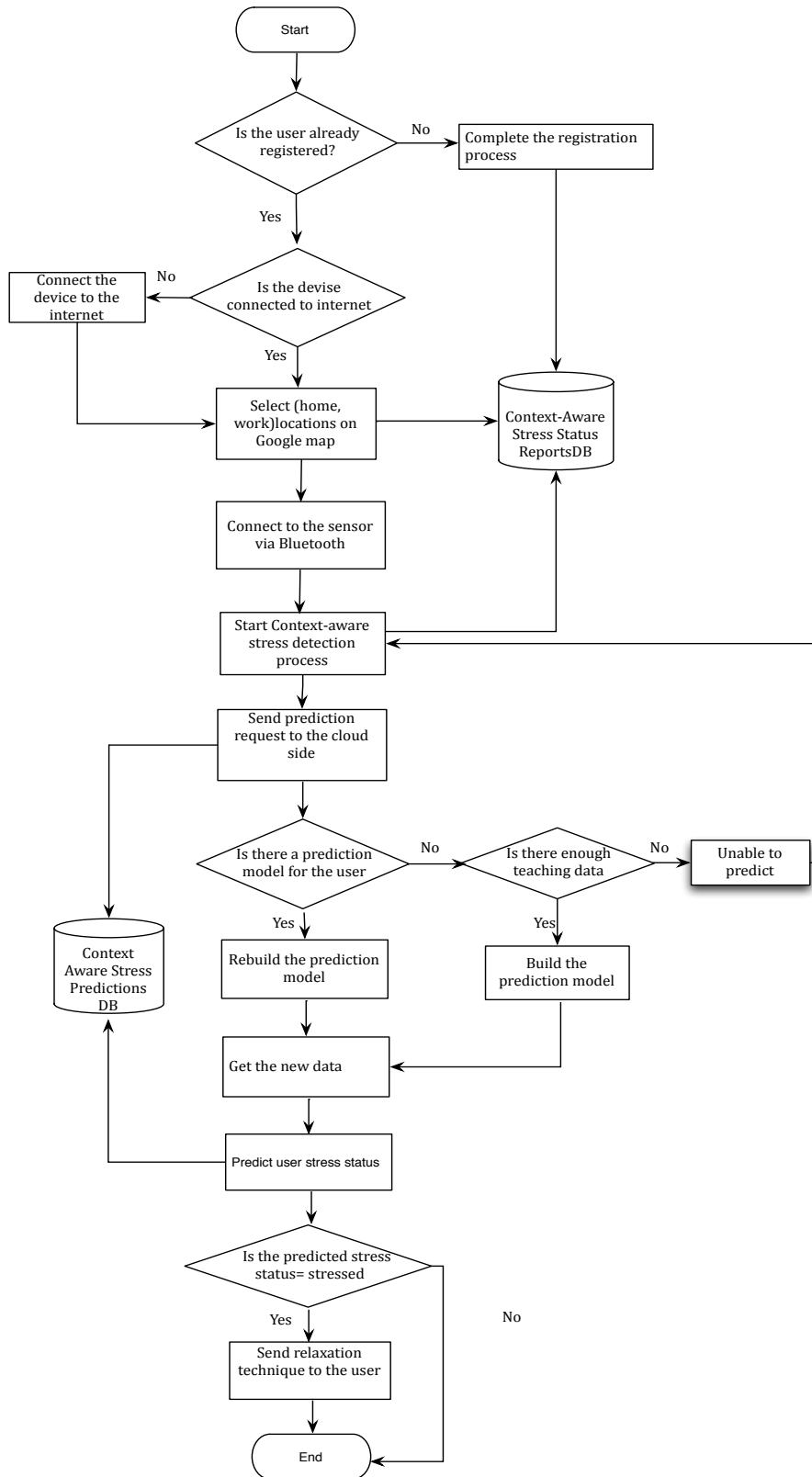


Figure 3.5: General CASP System Flowchart

The interactions between the system hardware and software components are illustrated with the aid of Unified Modeling Language (UML) diagrams, in Figures (3.6), (3.7) and (3.8).

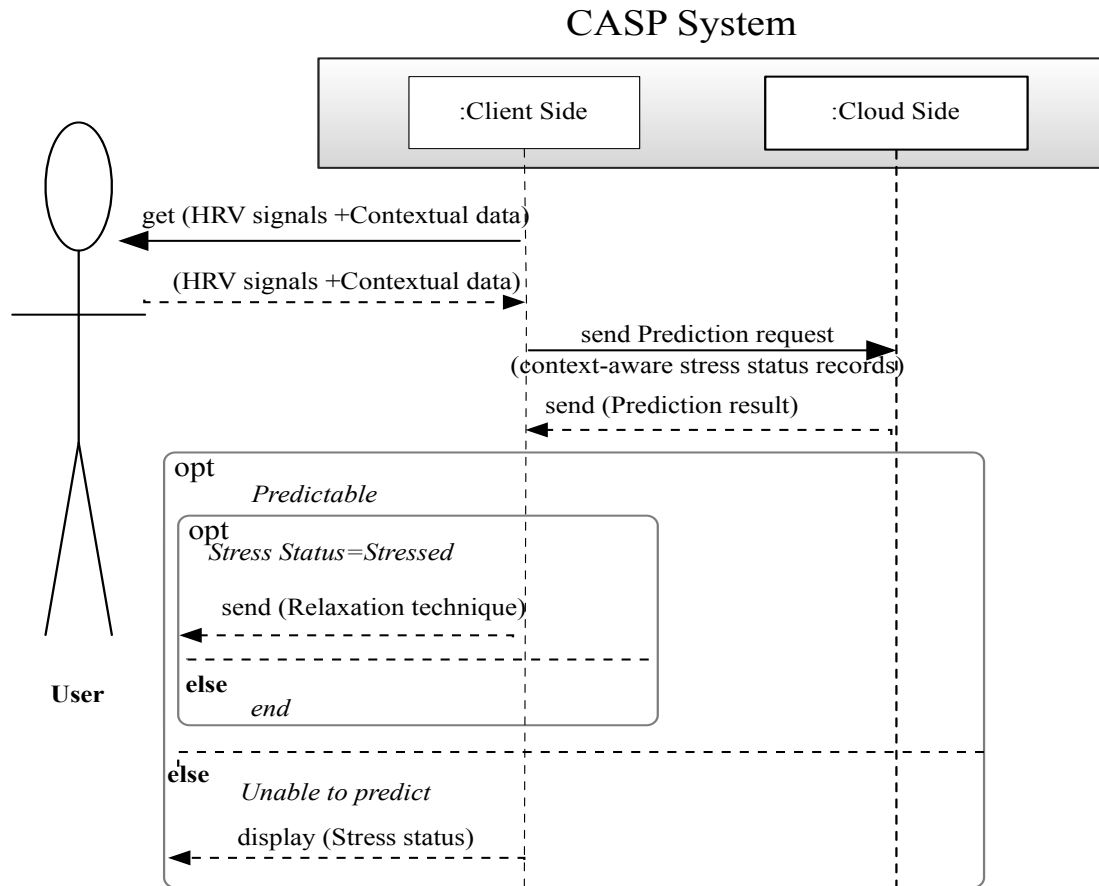


Figure 3.6: General CASP System Interaction Diagram

Figure (3.6) explains the interactions between the main components of the system for the client and the cloud sides. As shown in the diagram, the client side is responsible for collecting the physiological signals along with the contextual data from the user. In this component, the context-aware stress detection algorithm will generate the context-aware stress status reports of the users. These reports will then be sent to the cloud side as prediction requests. The result from the cloud side depends on the system's ability to predict. Either it is unable to predict, and in this case the detected stress status of the user will be displayed and a relaxation technique will be recommended for the user if a stressed status is detected. Or if it is able to predict, the predicted stress status will be sent; if the predicted status is stressed, a relaxation technique will be recommended to the user.

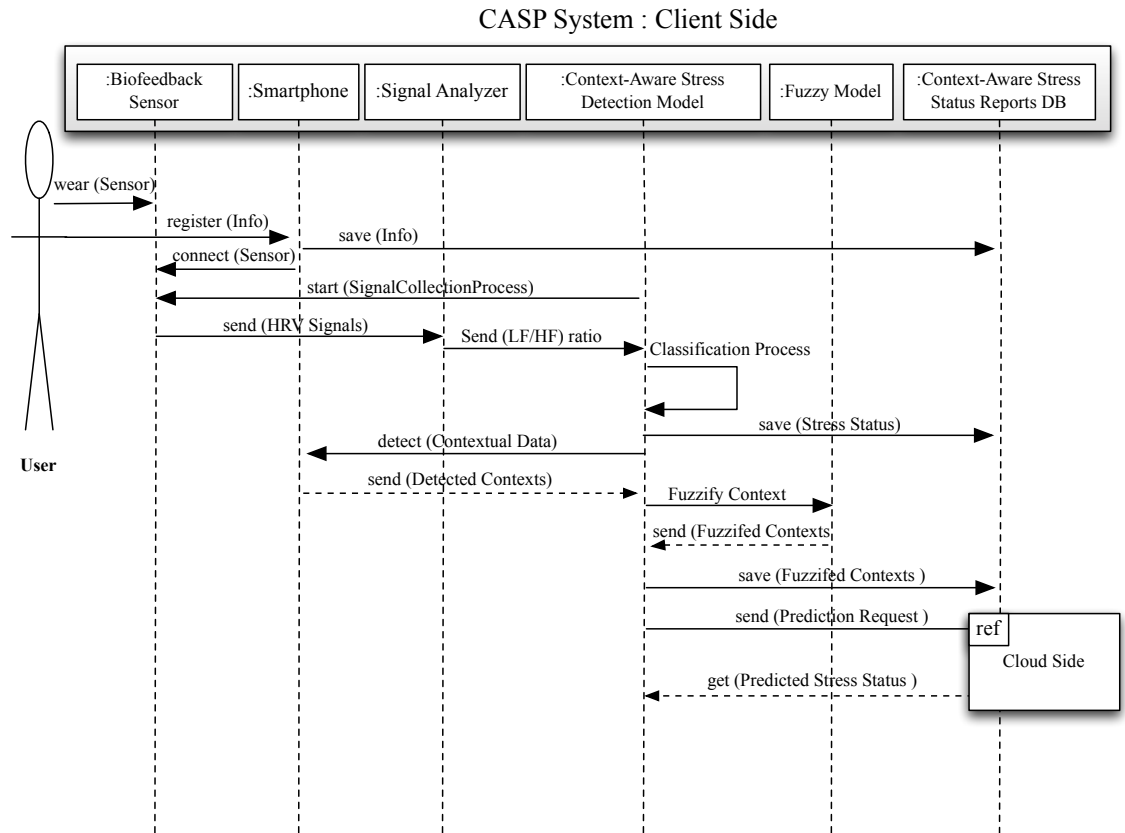


Figure 3.7: Client Side Interaction Diagram

Figure (3.7) provides a closer look at the interactions on the client side. It explains how the user starts by wearing a biofeedback sensor and completing the registration process in order to save all of their information in the local context-aware stress status reports database. The user then connects the sensor to the smartphone. After that, the context-aware stress status detection algorithm will start the process of collecting the signals, which will be sent from the sensor to the signal analyzer and then to the detection algorithm. In the algorithm, these signals will be classified into three different stress statuses. At the same time, the user's contextual data will be detected and fuzzified. As a result, a context-aware stress status record of the user will be saved in the database. Finally, this record will be sent to the cloud as a prediction request.

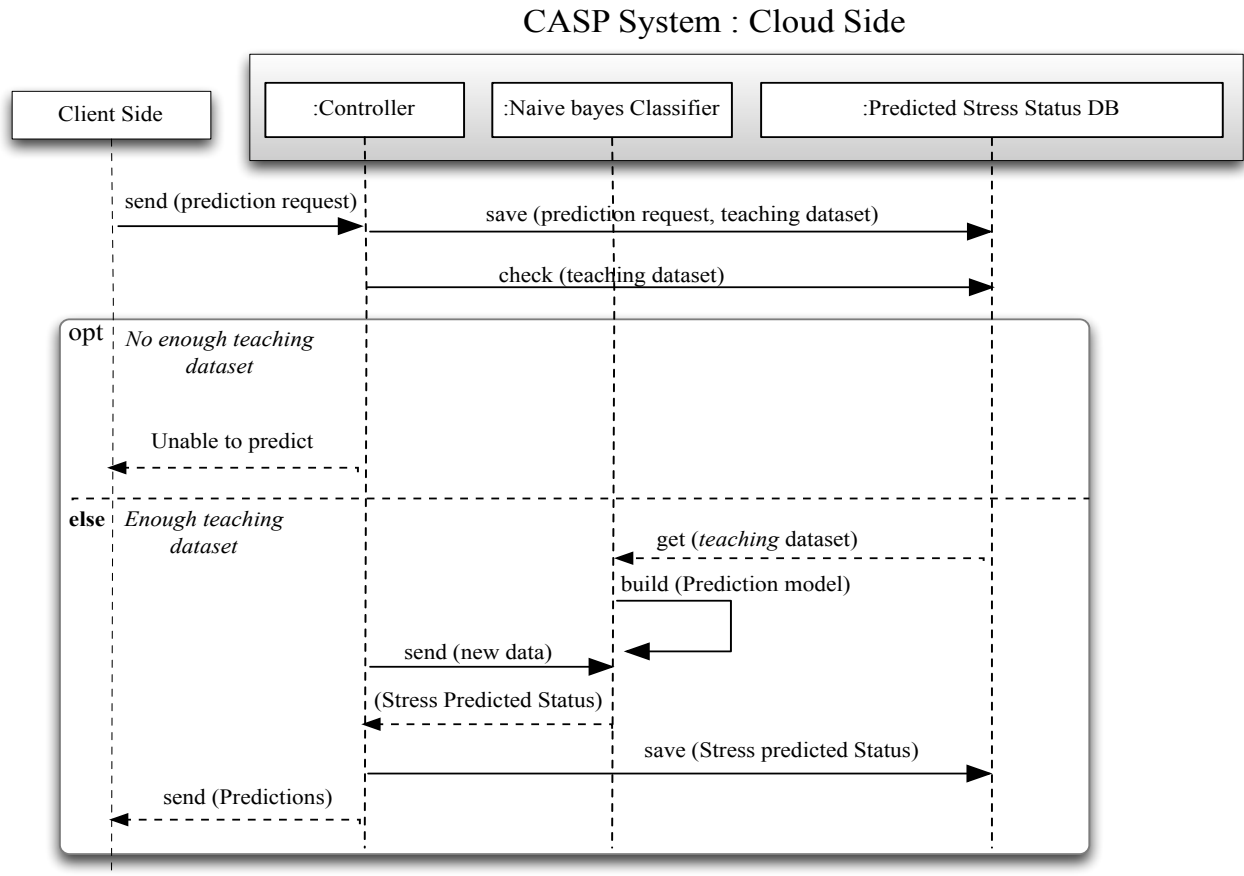


Figure 3.8: Cloud Side Interaction Diagram

The cloud side component interactions are shown in Figure (3.8). The controller in the cloud side will receive the prediction request sent from the client side and save it in the predicted stress status database. It will then verify its ability to predict the user’s stress status, which depends on the teaching dataset instances from which the model is learning. This step is accomplished by retrieving all of the user’s previous context-aware stress reports. The process leads to one of two options. The first one is when there are enough instances for the prediction to occur and the controller sends the retrieved teaching data from the database to the classifier to build the prediction model and do the predictions using the new data. Afterwards, each time a prediction request is received the prediction model will re-build using the teaching dataset. Finally, the predicted stress status of the user will be sent to the user and saved in the database for comparison and accuracy proposes. The second option is when there are not enough instances and the classifier is unable to predict the stress status. In this case, more instances are required.

Chapter 4

4 System Implementation

This chapter explains the hardware and software components of the system that we have implemented as a proof of concept. In Section 4.1 we provide the technical specifications for the hardware devices that are used. Explanation of the software components and the integrated application programming interfaces (APIs) are provided in Section 4.2. Finally, Section 4.3 provides all the system's graphical user interfaces.

The implemented system consists of the following main modules: Context-Aware Stress Status Detection and Context-Aware Stress Status Prediction, as shown in Figure (4.1).

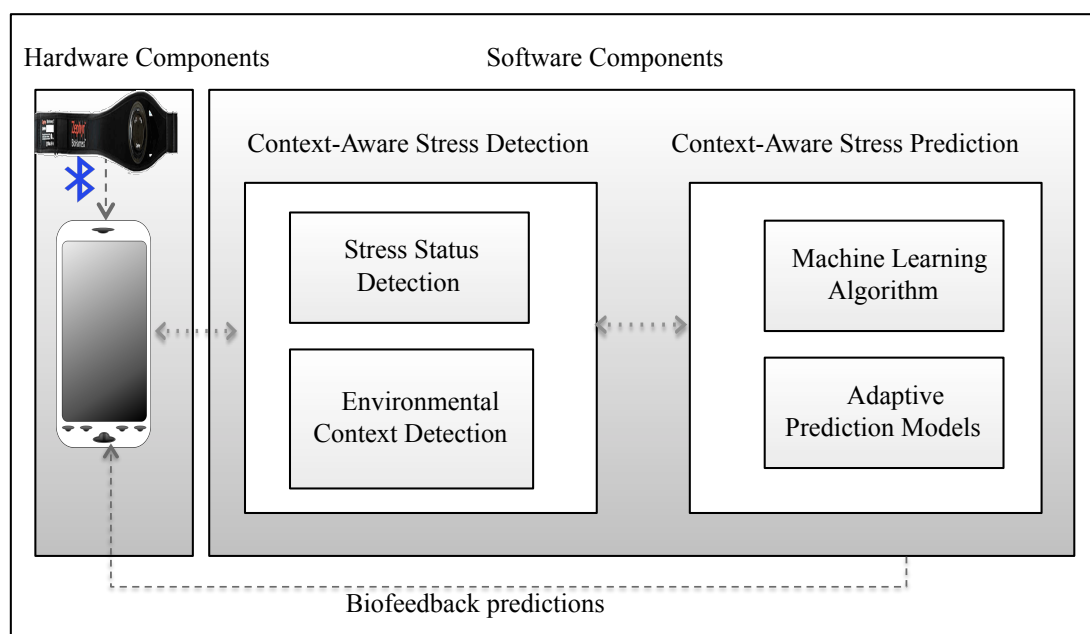


Figure 4.1: CASP System Software and Hardware Components

We implemented the system using the Android Studio Eclipse and platforms and Apache Tomcat, deployed in the Amazon EC2 cloud server. We also used the "qtuzzylogic" API for the context fuzzification process. We are using two different database engines; the first one is the SQLite Android local database, which is used to store the context-aware stress

status reports for each user. The second one is the PostgreSQL database, which is used on the cloud side to store the predicted stress status of all the users as it's more capable of handling the large number of users. Different APIs have also been incorporated into the system software to provide different functionalities. The APIs will be introduced below in each component where it is used.

4.1 Hardware Components

4.1.1 The Biofeedback Sensor

We used the Zephyr BioHarness belt, as shown in Figure (4.2), which is a wearable sensor that is tasked with collecting biological (ECG) signals. These signals show the electrical activity of the heart using two electrodes attached to the chest strap where the sensor is worn, on the user's torso. The relevant signals are sent to the targeted computing device via Bluetooth technology [51]. The context-aware stress detection component is responsible for running the Signal Processor and Signal Analyzer modules that process the received data.



Figure 4.2: Zephyr BioHarness

4.1.2 Android Device

The Android platform was selected based on its useful features and ease of implementation. This platform is a flexible open source that enables us to easily extend the source code. It also has the ability to access and manage the contextual data using the Android 'SensorManager'. Moreover, the 'context processing' has the required functions and software components to provide context data.

4.2 Software Components

4.2.1 Client Side Components

4.2.1.1 Signal Processor

The signal processor filters collect raw signals and extract the relevant information.

The LF\HF will be produced using Welch's method [50]. This method is used to estimate the signal's power spectral density (PSD) in order to calculate the frequency parameter, which is used in the stress management Android application in [2]. We use and follow the method used in [2] to conduct the analysis on five minutes worth of HRV data. The authors overlapped this period to produce time and frequency domain outputs every 30 seconds after collecting the benchmark value. The authors also proposed overlapping windows to take into consideration a signal's continuous nature.

4.2.1.2 Signal Analyzer

The signal analyzer is responsible for processing the collected signals and classifying the stress status based on the value of the LF\HF. This is done using the equations presented in Table (3.1) in Chapter 3.

4.2.1.3 Context-Aware Stress Status Detection

The stress monitoring activity starts once the benchmark values of the individuals are captured. The benchmark collection is a one-time process that should be performed during a neutral stress status, for a duration of five minutes. Further calculations are used to find the median and standard deviation of this value, which are then used in a comparison with each signal, produced in the signal processor and analyzed to get the classified stress status.

A. Context Data Detection

We detect five different contextual data in order to provide a context-aware stress status prediction service:

1. Weather conditions: we use the Open Weather Map API, which provides the current weather conditions of the city where the user is located. We obtain the responses of the API in JSON file format.
2. User locations: we retrieve the current location and the distance between the current location and the user's home and work. We do this using the Google Maps library provided in the SDK. Distances are calculated using the equations listed in Chapter 3.
3. Time of day: we get the time of day in Simple Date Format (HH:mm) using the *System.currentTimeMillis()* method that's always UTC, regardless of the system's time zone provided in the SDK.

4. Day of the week: we get the day of the week by using the *SimpleDateFormat* class provided in the SDK and the same *System.currentTimeMillis()* method.
5. Activity of the user: we get the user's activity by providing a list from which the user can select one of several activities. Each time a new session is started this list will appear.

B. Context Fuzzification

We use the qtfuzzylogic API to convert the discrete values of the detected context data into continuous data that has a degree of membership of more than one set. Each context consists of different sets, and each set has a range in which it describes the logical human reasoning limitations. Tables (4.1), (4.2), (4.3) and (4.4) below show the linguistic and numerical explanation of the fuzzified sets for all the contextual data. The following assumptions are used to study the gradual effect of the context on the stress status of the user. For example, Table (4.1) shows how the distance between the current location of the user and their home and work locations are converted to linguistic variables. Each variable has different ranges. For instance, the user is considered to be close to home or work if the distance is less than 12 km. The degree of membership is derived using the union membership function of the union of the two fuzzy sets for very close and close distances, as explained in the union set operation- Equation (3). We applied the same logic for all of the context data.

Table 4.1: Linguistic and Numerical Explanation of the Fuzzified Sets of the Distance Between the Current Location of the User and their Home and Work Locations.

Linguistic Explanation of the Distance	Numerical Values of the Distance (Km)		
	(a)	(b)	(c)
Very Close	1.000	6.660	12.700
Close	6.100	11.800	18.000
Far	12.500	18.200	24.200
Very Far	18.000	24.000	30.000

We take into consideration the different modes of transportation users may use (walking, biking, driving) when calculating the distance. Less than 1 km is considered to be very close, while more than 30 km is considered to be very far.

Table 4.2: Linguistic and Numerical Explanation of the Fuzzified Sets of Weather Conditions.

Linguistic Explanation of The Weather Conditions	Numerical Values of The Weather Conditions		
	(a)	(b)	(c)
Sunny-Clear	0.000	16.000	34.100
Partially Cloudy	24.000	41.000	60.000
Overcast clouds	46.000	63.000	80.000
Rainy	66.000	84.100	100.000

Table (4.2) shows the fuzzified weather conditions based on the cloud conditions and the chance of rain. At this stage we do not distinguish between rain and snow, as we didn't cover all the weather conditions.

Table 4.3: Linguistic and Numerical Explanation of the Fuzzified Sets of Days.

Linguistic Explanation of the Days	Numerical Values of the Days		
	(a)	(b)	(c)
Monday	0.000	14.180	29.040
Tuesday	22.320	38.180	53.040
Wednesday	46.320	60.500	77.040
Thursday	70.320	84.500	101.040
Friday	96.000	108.500	125.040
Saturday	118.320	132.500	149.040
Sunday	142.320	156.500	169.680

Table (4.3) shows the fuzzified day of the week, and we calculated the whole week based on the hours of each day, to analyze the gradual effects of the days on the user's stress status. The gradual affect of the day of the week on users stress status can be described

using fuzzy logic theory. It provides a valuable analysis of the causes of stress, where it specifies which day of the week contributes more to a specific stress status. A discrete day of the week for each measurement is also detected.

Table 4.4: Linguistic and Numerical Explanation of the Fuzzified Sets of Time.

Linguistic Explanation of the Time	Numerical Values of the Time in Hours/Minutes (HH:MM)		
	(a)	(b)	(c)
Morning	4:00 a.m.	8:00 a.m.	11:59 a.m.
Noon	-	12:00 p.m.	-
Afternoon	12:01 p.m.	15:00 p.m.	17:59 p.m.
Evening	17:00 p.m.	19:30 p.m.	20:30 p.m.
Night	20:00 p.m.	22:00 p.m.	23:59 p.m.
Midnight	-	00:00 a.m.	-
Deep Night	00:01 a.m.	2:00 a.m.	4:59 a.m.

Table (4.4) shows the fuzzified time of the day to allow us to study and analyze the gradual effect of the time on user's stress status. A discrete time is also detected for each stress status measurement. Where (a, b, c) are explained in the triangular membership function - Equation (2).

4.2.1.4 Context-Aware Stress Status Reports Database

We have selected the Android SQLite database (DB) to save the context-aware stress status data. This database has been selected because it meets our requirement of saving personal data locally for the context-aware stress status reports. The stress status will be updated every 30 seconds, along with the correlated contextual data, which will be continuously updated for any changes. The DB contains tables for the following:

- 1 User profile information.
- 2 Information about the individual's current location, home location, and work location.
- 3 The benchmark value.
- 4 The calculated median and standard deviation of the benchmark value.
- 5 The results obtained from the above calculations.

4.2.1.5 Database Tables

In this section we will provide a short description of each database entity. The entity relationship diagram of the CASP system database is presented in Figure (4.3).

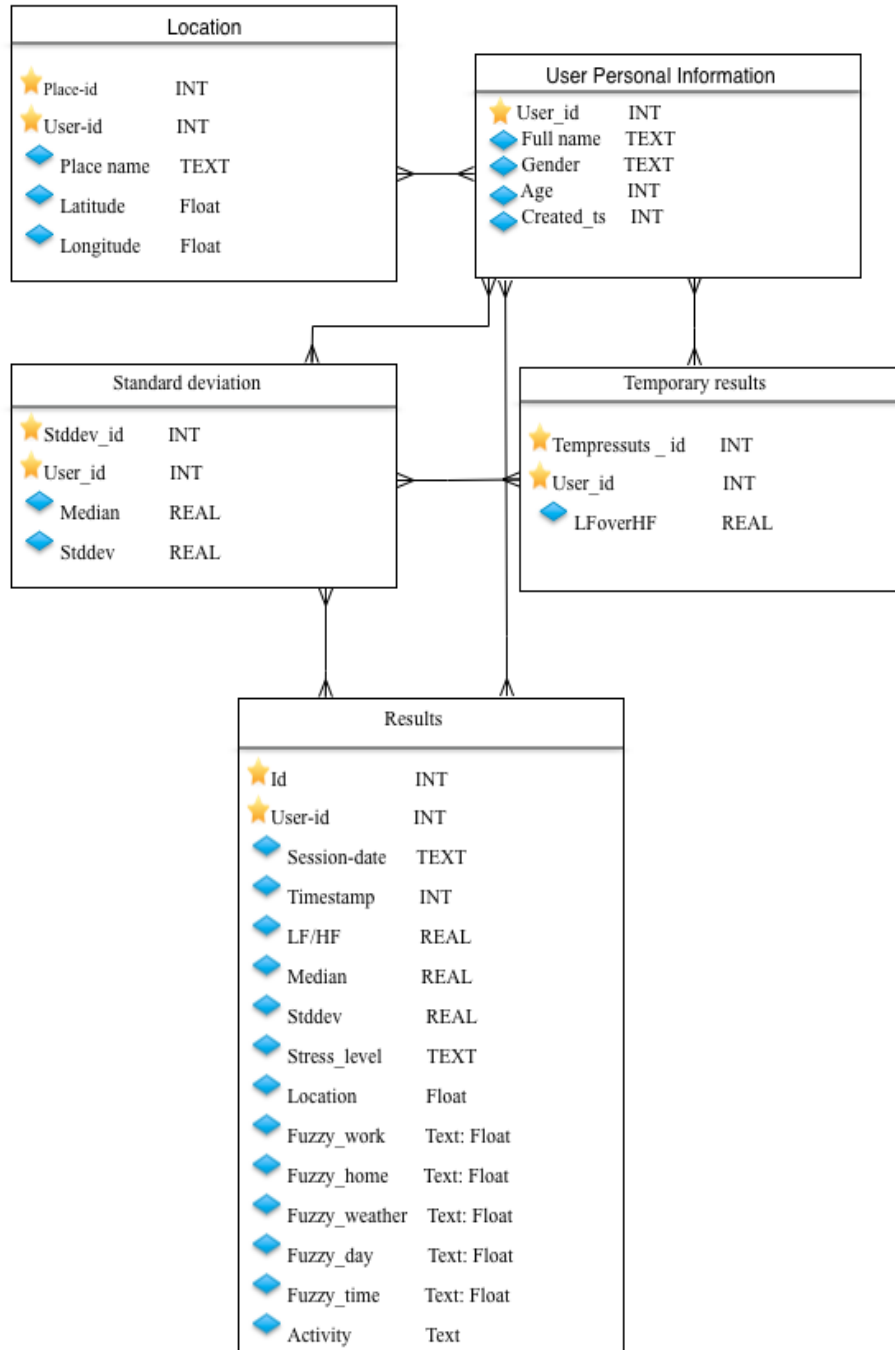


Figure 4.3: The Entity Relationship Diagram of the CASP System Database

A. Users

The User Personal Information database table shown in Table (4.5) holds the user's personal information and the time of each measurement.

Table 4.5: Users Information

Name	Type	Description
User_id	Integer	A unique identifier for each user
Full name	Text	The full name of each user
Gender	Text	The gender of each user
Age	Integer	The age of each user
Created_ts	Integer	The (year, day, hour, minutes, seconds) of each stress status measurement

B. Location

The Location database table shown in Table (4.6) holds all location information for each user. All places that have been reached by a user have a record in this table.

Table 4.6: Locations Information

Name	Type	Description
Place-id	Integer	A unique identifier for each place
User-id	Integer	A unique identifier for each user
Place-name	Text	A name for the place
Latitude	Float	A defined coordinate to specify a location using South–North position points
Longitude	Float	A defined coordinate to specify a location using East–West position points

C. Users Benchmark Values

The users benchmark values database table shown in Table (4.7) holds all parameters required to collect the benchmark value for each user.

Table 4.7: Users Benchmark Values

Name	Type	Description
Tempressuts_id	Integer	A unique identifier for LF\HF signals to collect the benchmark
User_id	Integer	A unique identifier for each user
LFoverHF	Real	The measured signals of the LF/HF ratio

D. Standard Deviation and Median Values

The Standard Deviation database table shown in Table (4.8) holds all parameters and required calculations of medians and standard deviations on the collected benchmark values from the Users Benchmark Values table to classify the stress levels.

Table 4.8: Standard Deviation and Median Values

Name	Type	Description
Stddev_id	Integer	A unique identifier for each standard deviation has been calculated
User_id	Integer	A unique identifier for each user
Median	Real	The calculated median of the benchmarks
Stddev	Real	The calculated standard deviation of the benchmarks

E. Results

The Results database table shown in Table (4.9) holds all the calculated context-aware stress statuses. Every 30 seconds a collected signal is saved with all of the contextual data that describes the surrounding environment at that time.

Table 4.9: Context-Aware Stress Status Results

Name	Type	Description
Id	Integer	A unique identifier for each executed function for each user
User-id	Integer	A unique identifier for each user
Session-date	Text	The date of each session
Timestamp	Integer	The year, day, hour, minutes, and seconds of each measurement when the data was sampled
LF/HF	Real	The measured signals of the LF/HF ratio
Median	Real	The calculated median of the five minute collection of the LF/HF signals
Stddev	Real	The calculated standard deviation of the five minute collection of the LF/HF signals
Stress_level	Text	The calculated stress level
Location	Float	The user's current location
Fuzzy_work	Text: Float	The calculated and fuzzified distance between the user's current location their work
Fuzzy_home	Text: Float	The calculated and fuzzified distance between the user's current location and their home
Fuzzy_weather	Text: Float	The fuzzified weather conditions in the user's current location (city)
Fuzzy_day	Text: Float	The fuzzified day in which the measurements are taken
Fuzzy_time	Text: Float	The fuzzified time in which the measurements are taken
Activity	Text	The current activity of the user

4.2.2 Cloud Side Components

We are using the Apache Tomcat 8 platform, the Servlet container, and Java 8 in the Amazon EC2 cloud server. The system communicates with the web server application via API using an XML format.

4.2.2.1 The Controller

The controller is responsible for dataset preparation, which arranges data coming from XML files (Prediction requests) and from the PostgreSQL DB teaching dataset. It then checks if the number of instances stored in the database is enough to build the prediction model or not. The prediction models will be generated using the Weka API that builds adaptive\personalized prediction models using all historical data.

4.2.2.2 Predicted Stress Status Database

We are using the PostgreSQL database, which has the ability to save a lot of data from a large number of users. In this database we have the history data as well as predicted stress status shown in Table (4.10). This table holds all the predictions that have been calculated and sent to the user. The table consists mainly of the sensed stress level along with all of the contextual data. The teaching dataset will be retrieved from this database in order to build the prediction models.

Table 4.10: The Context-Aware Stress Status Prediction

Name	Type	Description
Id	Numeric	A unique identifier for each measurement
StressLevel	Text (string)	The calculated\sensed stress level
PredictedStress	Text	The predicted stress level
LocationLat	Real (integer)	A defined coordinate to specify a location using South-North position points
LocationLon	Real	A defined coordinate to specify a location using East-West position points
WorkText	Text	Distance between the user's current location and their work
WorkValue	Real	
HomeText	Text	Distance between the user's current location and their home
HomeValue	Real	
WeatherValue	Real	Weather conditions in the user's current (city)
DayValue	Real	The day of the week in which the measurements are taken

TimeValue	Real	The time of the day at which the measurements are taken
Activity	Text	The current activity of the user
UserID	Integer	A unique identifier for each user
Session	Text	The date of the session

4.3 CASP System Interfaces

The graphical user interfaces of the system are summarised in Figure (4.4).

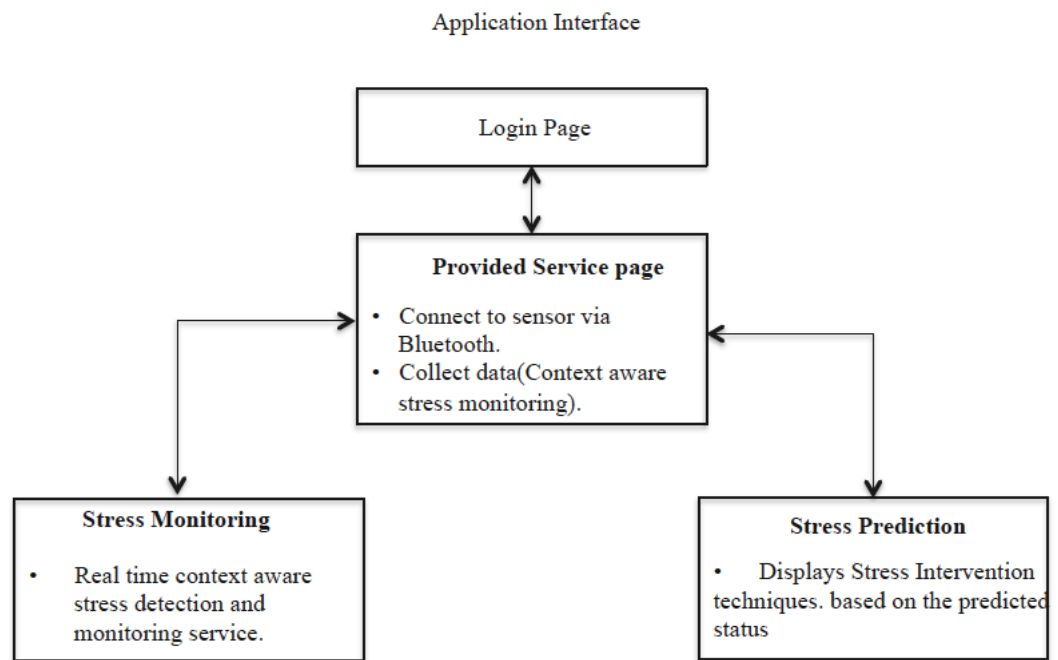
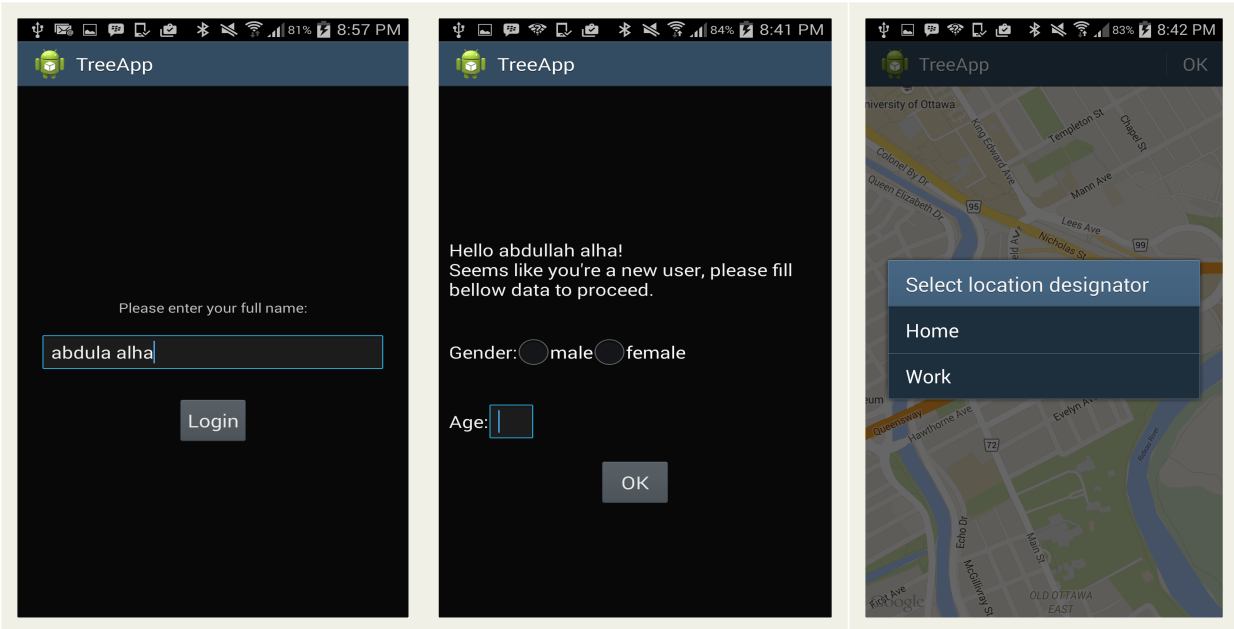


Figure 4.4: CASP System Interfaces Architecture

4.3.1 Login Interface

Each user must log in for the first time in order to save their required personal information as shown in Figure (4.5) (a).



(a)

(b)

(c)

Figure 4.5: CASP System Interfaces

4.3.2 Personal Information Interface

The user then need to enter additional information such as gender and age, as Figure (4.5) (b) shows.

4.3.3 Locating Related Places Interface

The user then has to provide their home and work locations, as shown in Figure (4.5) (c), in order to study their effects on the stress status, and also the effect of being away or getting closer to one of them.

4.3.4 Stress Monitoring Interface

After starting the connection between the system and the biofeedback sensor via Bluetooth, the system will then show a message saying “*The system is learning about you*”. The CASP system requires this step in order to collect the benchmark values for 5 minutes. These benchmark values will be used in the calculations that classify the stress level using the context-aware stress detection algorithm. The stress monitoring page shows the user's monitored stress level represented as a tree and flowers. The tree gets greener

and shows growing flowers whenever the user is relaxed, and the opposite occurs with a stressed status. We are using stress status presentation presented in [2].

4.3.5 Stress Intervention Interface

Based on the predicted stress status, if the user's stress status is "stressed", a video will be presented as a recommended relaxation technique [17], as shown in Figure (4.6).

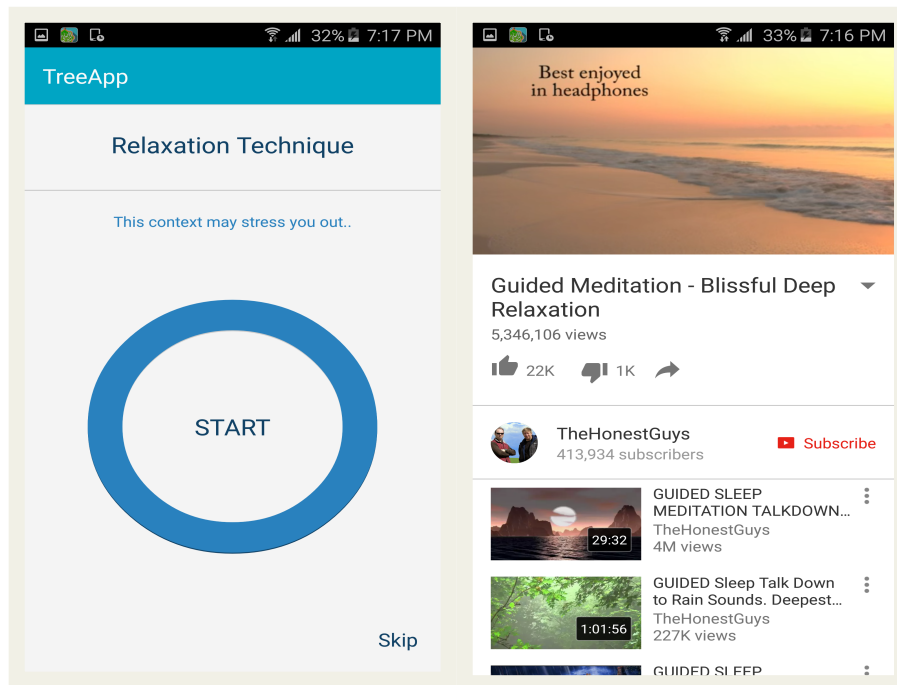


Figure 4.6: Stress Intervention Interface

Chapter 5

5 System Evaluation and Experimental Results

In this chapter, we present the evaluation of the CASP system's performance, which answers the main research question of whether the deployment of a user's contextual information could contribute to the prediction of their stress status. We evaluate the prediction performance of the selected classifier compared to the ground truth of the sensory data using well-known evaluation metrics. Moreover, the evaluation process includes the assessment of the personalization and the adaptation features of the prediction models. With the aid of graphical representations, we will show how each context contributes to changing the user's stress status in a subjective way. The adaptability of the system will be evaluated by assessing the ability of the system to rebuild the prediction models in the cloud using the new contextual data.

The chapter starts with the evaluation hypothesis in Section 5.1. The experimental setup follows in Section 5.2. Finally, in Section 5.3, the experimental results of the prediction performance will be presented.

5.1 Evaluation Hypothesis

1. If we are able to provide a personalized stress measurement method that meets the subjective nature of the stress, we will be able to generate a personalized prediction system as well.
2. If we are able to find the correlation between the contextual data and the stress status, we will be able to consider some of the context data as stressors.
3. If we are able to re-build the prediction models to adapt to the changing nature of the contextual data, we will be able to provide an adaptive prediction system.
4. If we are able to predict the user's stress status, we will be able to send stress interventions in the early stage to prevent the occurrence of stress.

All the hypotheses mentioned above are alternative. The superiority of the selected algorithm to other algorithms will not be considered in this evaluation because it has been selected based on specified criteria as explained in detail in Chapter 3.

5.2 Experimental Setup

The main goals of this experiment are to assess the validity of our system and to verify its ability to predict the individual's stress status (*dependent variable*) based on the current contextual information (*independent variables*).

5.2.1 Participants

We have conducted a real life non-controlled experiment on 5 subjects: 3 females and 2 males between the ages of 20 and 30. None of the participants have a heart disease, hypertension, or other diseases that may affect the stress monitoring process. The detailed description of the subjects is provided in Table (5.1). The experiment was conducted on the participant's real life regular schedules. Participants were asked to give a brief description of their daily activity (general information about their life) during the measurement period.

The description provided by the participants will be presented in the following paragraphs. Participant 1 was working on a research paper for most of the time, while also programming and writing documents. The participant mentioned the task overload that she was trying to handle at that specific period in time. Participant 2 was in the process of writing a master's thesis and was preparing for their master's defense. This participant was also taking driving lessons. In addition, she talked about her desire to lose weight and start a diet program again. Participant 3 was designing and conducting a research experiment as well as programming. She had a very busy schedule. Participants 4 and 5 were both in the middle of a vacation. Participant 4 traveled to another country, and was working to develop some new personal skills. The user mentioned the sleep problems he was facing lately due to the traveling. Participant 5 was enjoying time with family most of the time with no important tasks at all. This participant did face certain emotional events during the time of the experiment.

These descriptions provide us with an overview of the most important events and activities in which the participants were engaged. They gave us a better understanding of the emotional waves that are shown in the experimental results. A summary of the participant discretions is provided in Table (5.1).

Table 5.1: Detailed Description of the Subjects

Subject Number	Gender	Age	Occupation	Brief Description
1	Female	28	Graduate Student Computer Science	– Performing research and programming
2	Female	27	Graduate Student Biology	– Spent most of the time writing a Master’s thesis – Preparing for the Master’s defense
3	Female	25	Graduate Student System Science	– Programming and conducting research experiments
4	Male	23	Undergraduate Student: Biotechnology	– On vacation – No organized schedule – Actively learning new skills
5	Male	21	Undergraduate Student: Radiology	– On vacation – Being socially active – Engaging in fun activities with family

5.2.2 Experimental Phases

The experiment was divided into four phases:

1. Data collection phase
2. Model training and testing phase
3. Building and re-building the prediction models phase
4. Intervention phase

Firstly, when the CASP system started, the participants were asked to provide their personal information (name, age, gender, home and work locations) in order to create their personalized profile for the context-aware stress status records. They were then asked to wear the biofeedback belt and connect it with the system via Bluetooth by clicking on the “connect to the biofeedback belt” button. They were then asked to select the activity that they were currently performing from a list of activities. By doing this, the context-aware stress monitoring process was started.

A detailed description of the phases is provided below:

Phase 1: Data Collection

This phase starts with the collection of benchmarks for further use in the stress assessment process using the LF\HF parameter and the equations presented in table (3.1). This process takes five minutes. After collecting the benchmarks, the system starts the process of recognizing a user's stress status, with a high sensitivity to the small stress changes as explained in table (3.1). The stress status is recognized along with the user's current contextual data during their regular scheduled activities. The duration of the data collection phase for each participant which takes (3-5) days , and the number of instances of the personalized datasets are shown in Table (5.2).

Table 5.2: Participant Datasets

User ID	Dataset Instances	Duration of Use	
		Days	Hours
User 1	806	3	7.15
User 2	2130	5	18.20
User 3	498	3	4.15
User 4	863	3	7.19
User 5	1264	3	10.50

Each dataset has seven different attributes (the contextual data) for three different classes (the stress statuses). The data type of these attributes is mixed (continuous, discrete). More information about the attributes is provided in Table (5.3).

Table 5.3 Attributes Description

Attributes	The Discrete Values	The Continuous Values Range
Time of day	Morning, Noon, Afternoon, Evening, Night, Midnight, Overnight	(0.00, 24.00) (HH:MM)
Day of week	From: Monday/Sunday	(0.00, 168.00)
Current location	Not applicable	Location's latitude and longitude
Distance from home	Very close, close, medium, far	(0.00, 30.00) (Km)
Distance from work	Very close, close, medium, far	(0.00, 30.00) (Km)
Current weather (based on the chance of rain)	Sunny, partially cloudy, overcast, rainy	(0.00, 100.00)
Activities	Studying, eating, relaxing, walking, meeting, biking	Not applicable

Phase 2: Training and Testing

For training and testing we used the 10-fold cross validation by dividing each dataset into 10 parts and using 9 parts for training and the last part for testing. This process is repeated 10 times, and each time a different segment is used for the testing. The 10 results are averaged as a result of the 10-field cross-validation.

Phase 3:

A. Building the Prediction Models

In this phase, the prediction model is built in the cloud using the previously collected data. Building the prediction model takes less than 20 seconds. The current contextual data will be used as an unseen sample to predict the user's current stress status in less than 2 seconds. The predicted status is saved in the database and sent back to the mobile application. In a case where there are not enough teaching instances to build the model, the prediction fails and more instances are required.

B. Re-building the Prediction Models

Based on the rapidly changing nature of stress and environmental contexts, as well as their relation to the states/context dependent memory of the user, we provided a service that adapts to these changes to increase the performance of the prediction process. Therefore, the last predicted status is compared with the actual sensed stress status. It is used along with the previously collected data in the process of re-building the model in order to generate the new predictions in less than 22 seconds. The prediction models will be re-built each time a prediction request is sent with new instances based on the availability of the Internet connection.

Phase 4: The Early Intervention

In the case where the prediction result is a “stressed” status, a relaxation technique will be suggested to the user as soon as the system has captured possible stressors in the current contextual data.

5.2.3 Apparatus

- We use the Zephyr BioHarness biofeedback sensor, which measures the user’s stress levels when linked to a smartphone via Bluetooth [51].
- Android smartphone S4. A full description of the used apparatus is provided in Chapter 4.

5.3 Evaluation of the Experimental Results

We use two different methods for the evaluation of the system:

- The cross validation method to assess the performance of the Naive Bayes classifier to deliver accurate personalized predictions.
- A statistical analysis that shows the difference between the actual and the predicted stress statuses.

5.3.1 Evaluation Metrics

We use the confusion matrix of a 10-fold cross validation. Consider the Naive Bayes algorithm running on the individual dataset in WEKA. We obtained three classes for a 3x3 confusion matrix for each user’s dataset. Based on the confusion matrix we obtained the following well-known evaluation measurements: accuracy, precision, recall,

and F-measure. Additionally, the predicted vs. actual (which are the basic measures), are used to evaluate the prediction results.

Recall is the measurement that evaluates the predicted status based on only the relevant status in the database, as shown in Formula (11). Precision, on the other hand, is evaluated based on both the total number of irrelevant and relevant status predictions, as shown in Formula (12). Both precision and recall are calculated using each of the TP-FP (13), (14) rates from the confusion matrix. In another words, precision is known as the positive predictive value while recall shows the sensitivity. Their trade-off is the F score (15).

$$Recall = \frac{|{\textit{relevant Stress Status}} \cap {\textit{predicted Stress Status}}|}{|{\textit{relevant Stress Status}}|} \quad [11]$$

$$Precision = \frac{|{\textit{relevant Stress Status}} \cap {\textit{predicted Stress Status}}|}{|{\textit{predicted Stress Status}}|} \quad [12]$$

TP Rate: when the predicted stress status is the same as the actual stress status.

FP Rate: when the predicted stress status is not the same as the actual stress status.

$$F = 2 \cdot \frac{Precision \times recall}{Precision + recall} \quad [13]$$

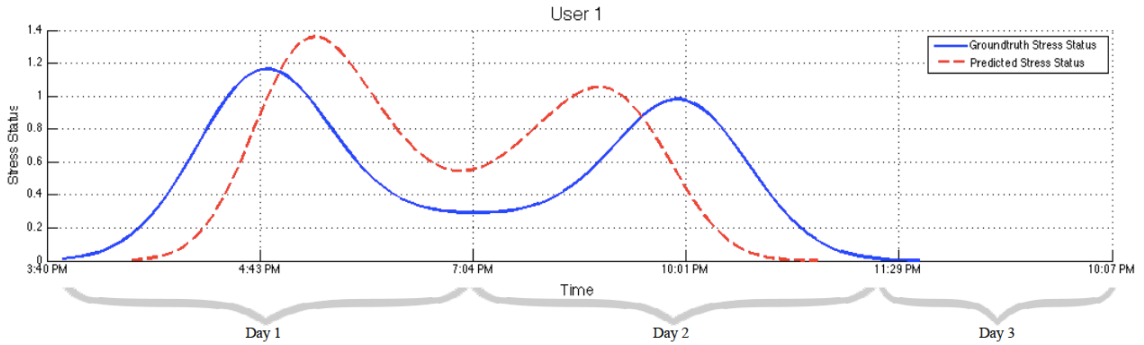
F-Measure: The trade-off of precision and recall

Table 5.4: Prediction Model Accuracy

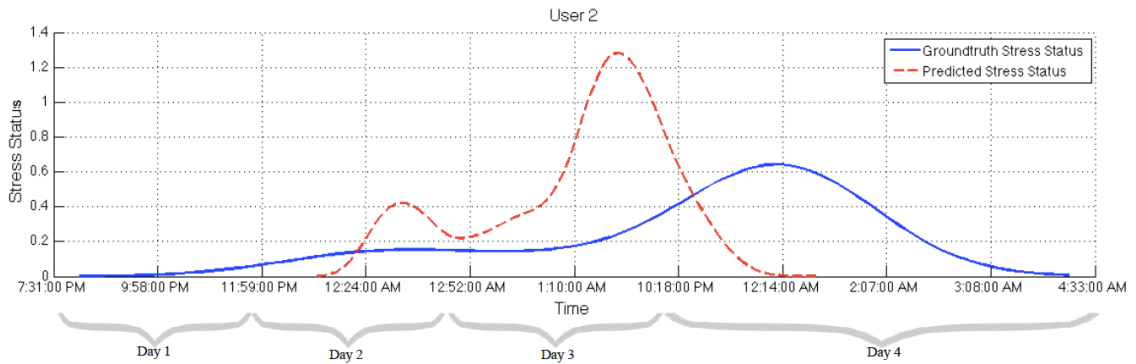
User ID	Correctly Classified Instances
1	84.9876%
2	56.0094%
3	88.755%
4	86.7903%
5	75.0791%
Average	78.32%

All of the accuracies presented in Table (5.4) were obtained by averaging the results from 10 runs of 10-fold cross-validation. The table shows that all of the

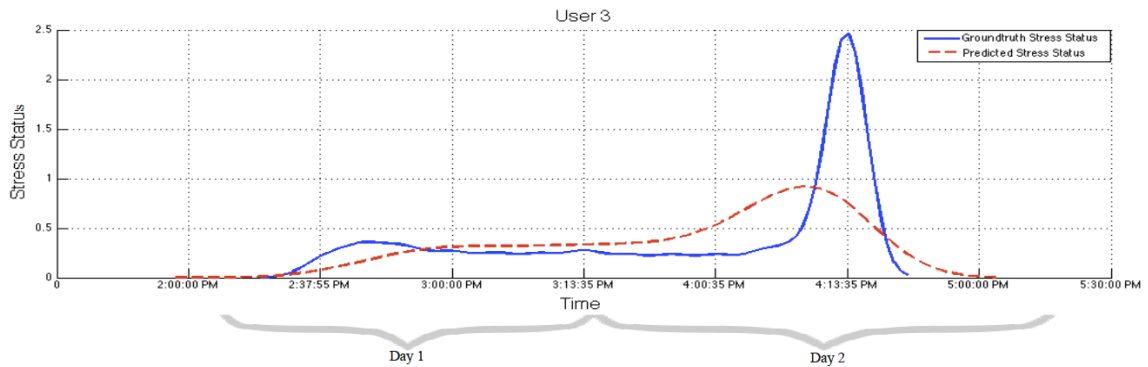
personalized models produced acceptable prediction accuracies, with 78.32% as the average. Figure (5.1) shows a comparison between the predicted stress status obtained from the classifier and the ground truth stress status (sensory data) over time for all the five users.



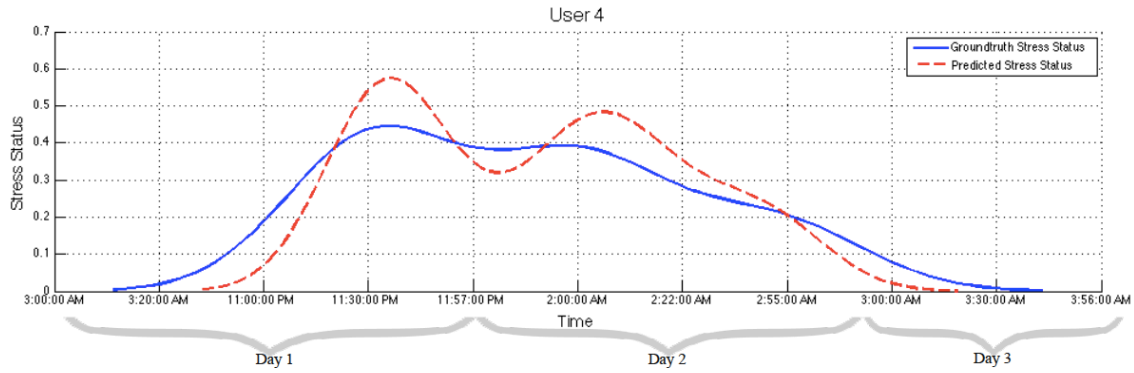
A: Comparison between ground truth stress status and predicted stress status for user 1



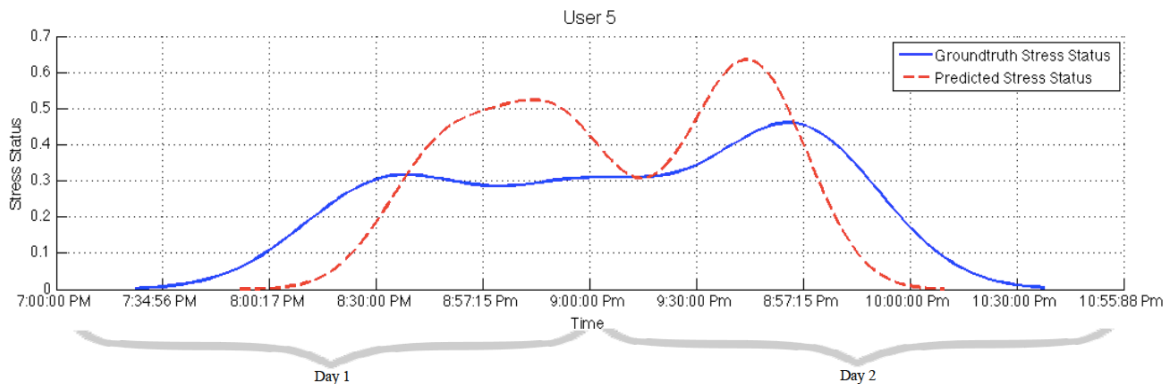
B : Comparison between ground truth stress status and predicted stress status for user 2



C: Comparison between ground truth stress status and predicted stress status for user 3



D: Comparison between ground truth stress status and predicted stress status for user 4



E: Comparison between ground truth stress status and predicted stress status for user 5

Figure 5.1: Comparison Between Ground Truth Stress Status and Predicted Stress Status for All Users

The red curves represent the predicted status, which means the performance of the classifier in predicting the user’s stress status. The ground truth of the sensed stress status is represented by the blue curves. The X-axis in the figures represents the day/time of the stress status. The Y-axis shows the stress status level, ranging from 1 to 3. An increase represents higher stress levels. We made the following observations on the overall performance of the classifier: The red curves almost fit the blue curves with some rightwards shift. The red curves display similar responses to the blue curves (showing the status of the user’s stress). The classifier promotes better performance in the cases where the prediction model trained with more instances, as shown in Figure (5.1) (D) and (E). However, with less training data the classifier still shows good performance as displayed in Figures (5.1) (A) and (C).

Table 5.5: Detailed Accuracy by Class for Individual Prediction Models using Precision, Recall, and F-Measure Parameters.

User ID	Class	Parameters		
		Precision	Recall	F-Measure
1	NEUTRAL	0.847	0.884	0.865
	STRESSED	0.857	0.807	0.831
	RELAXED	0.6	1	0.75
	<i>Average</i>	0.851	0.85	0.85
	<i>STDEV</i>	0.146	0.097	0.059
2	NEUTRAL	0.202	0.607	0.313
	STRESSED	0.418	0.48	0.559
	RELAXED	0.996	0.48	0.648
	<i>Average</i>	0.829	0.56	0.603
	<i>STDEV</i>	0.411	0.073	0.174
3	NEUTRAL	0.8	0.846	0.822
	STRESSED	0.714	0.769	0.741
	RELAXED	0.945	0.917	0.931
	<i>Average</i>	0.891	0.888	0.889
	<i>STDEV</i>	0.117	0.074	0.095
4	NEUTRAL	0.946	0.912	0.929
	STRESSED	0.816	0.816	0.816
	RELAXED	0.798	0.859	0.827
	<i>Average</i>	0.87	0.686	0.869
	<i>STDEV</i>	0.081	0.048	0.062
5	NEUTRAL	0.998	0.769	0.868
	STRESSED	0.491	0.836	0.619
	RELAXED	0.838	0.659	0.738
	<i>Average</i>	0.826	0.751	0.768
	<i>STDEV</i>	0.259	0.089	0.125

In Table (5.5), a closer look at the detailed accuracy for each class is provided. This data will be analyzed using precision, recall, and F-measure. Using the recall values, which are calculated using (11), we achieve different recall results for the personalized models. As shown in Table (5.5), the highest value is 0.888 and the lowest value is 0.56. As we want to identify as much relevant context data related to a specific stress status as we can, the lowest value we achieved is considered to be acceptable compared with a perfect recall of 1. On the other hand, the precision values are calculated using (12) and their values are above 0.8. The averaged values of the F-measure are also provided in Table (5.5), where we see that most of the models achieved 0.8. These values indicate a balanced mean between the precision and recall values.

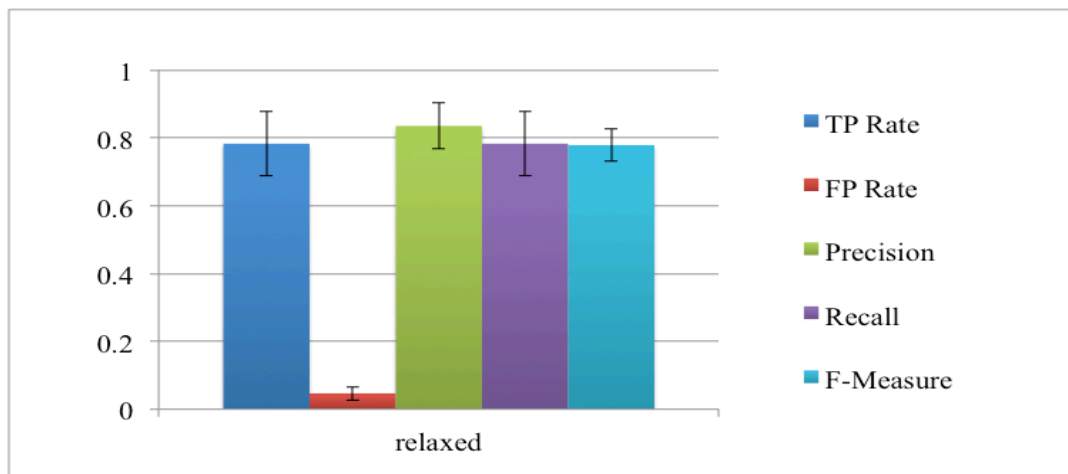


Figure 5.2: The prediction Results Classified by the Relaxed Status for All Users

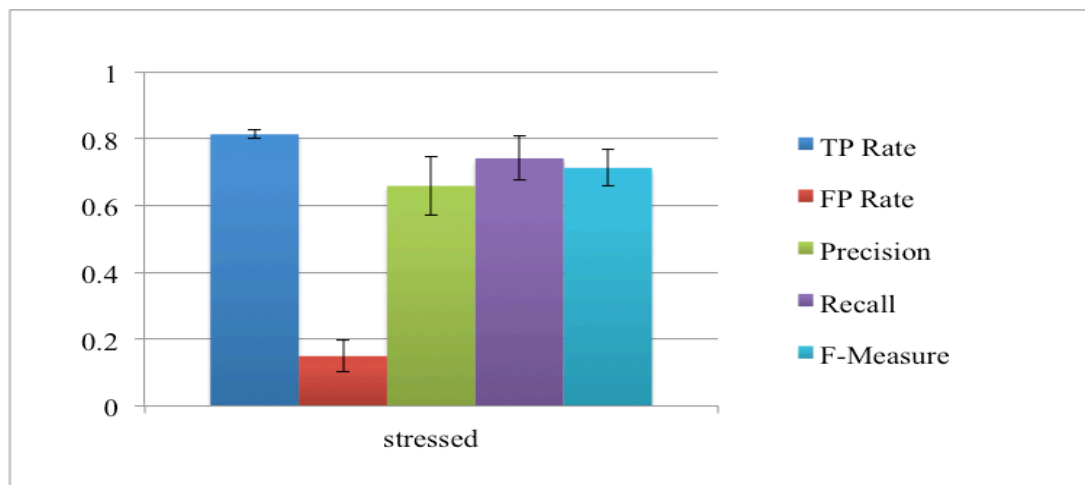


Figure 5.3: The Prediction Results Classified by the Stressed Status for All Users

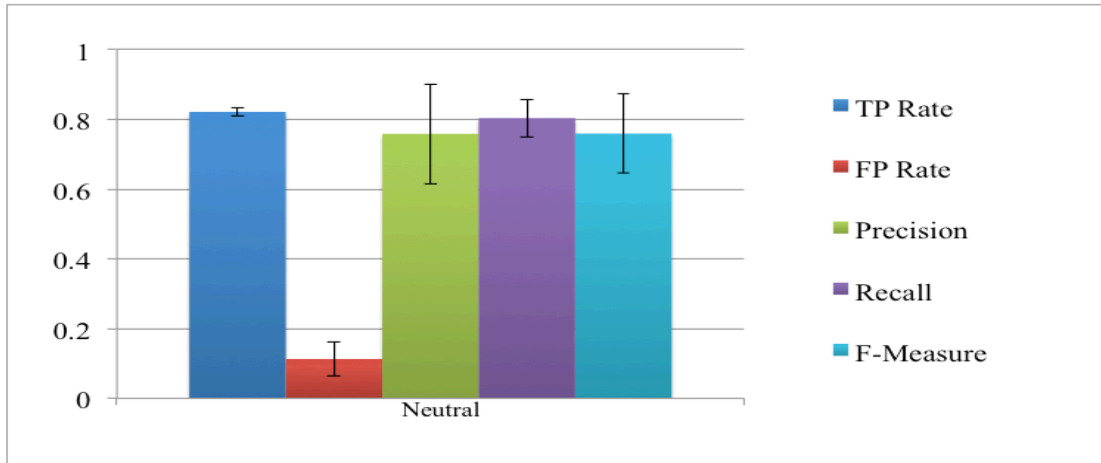


Figure 5.4: The Prediction Results Classified by the Neutral Status for All Users

The prediction results classified by the stress status for all users are shown in Figures (5.2), (5.3) and (5.4). There are five main measurements that can be advanced to analyze the prediction results in each class. The False Positive rate indicates the incorrectly classified stress status, shown in all three graphs to be less than 0.2. The rate of the True Positive exceeds 0.8, which implies a high amount of correctly classified stress statuses achieved among the three classes. The F-measure exceeds 0.6 for the three stress status classes. This measure summarizes the values of the precision and of the recall by calculating the mean of these two values. The closer these values get to 1, the more perfectly they meet the classification requirements.

As an example, we show the analysis of the users data below. The graphs in Figures (5.5), (5.6), (5.7), (5.8) and (5.9) illustrate how much each context acts as a stressor for each user. Graphs (a), (b), and (c) show the percentage of stress experienced doing each activity, each day of the week, and at each time of day. Graphs (d) and (e) show the different stress statuses of the users at their home and work locations for the first three users and at the home location of the last two users, as they were on vacation during the experiment.

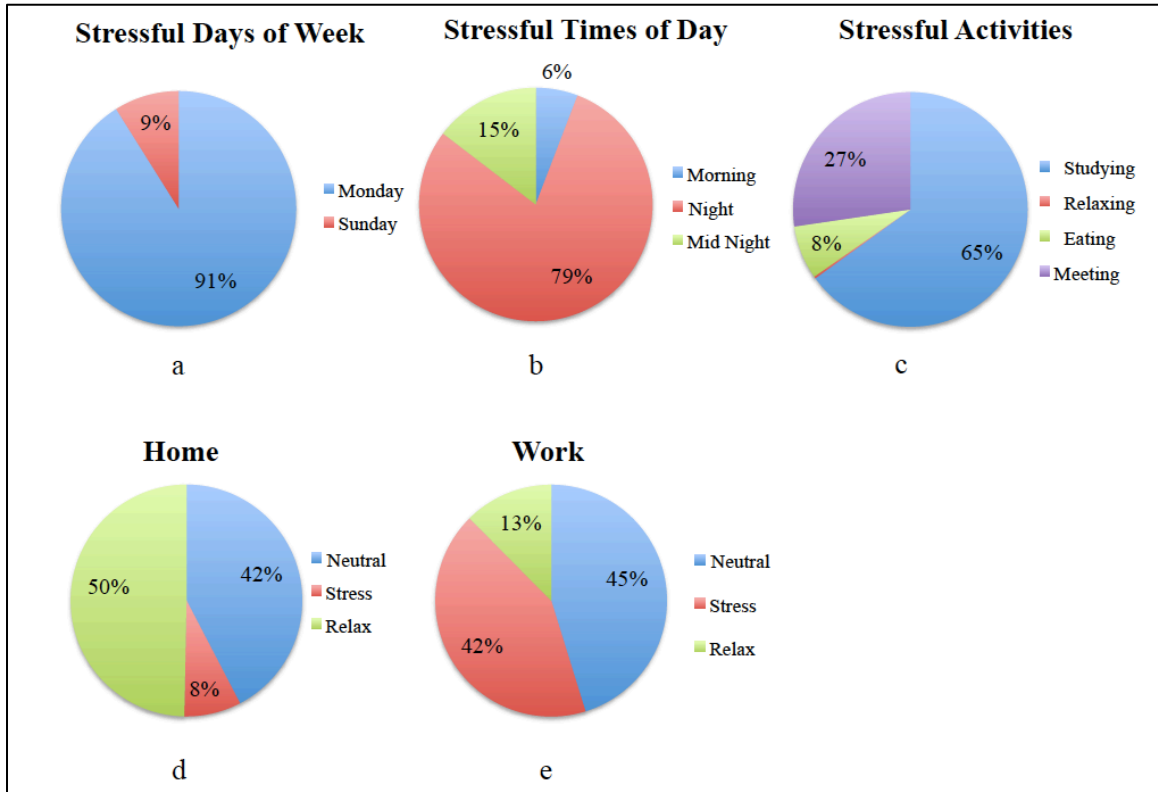


Figure 5.5: Analysis of the Way Contextual Data Affects the Stress Status of User One. A represents the stressful days of the week. B represents the stressful times of the day. C represents the stressful activities. D represents the different stress statuses at home. E represents the different stress statuses at work.

As shown in Figure (5.5), studying and meeting are the most stressful activities, Monday is the most stressful day of the week, and night is the most stressful time of the day for user one.

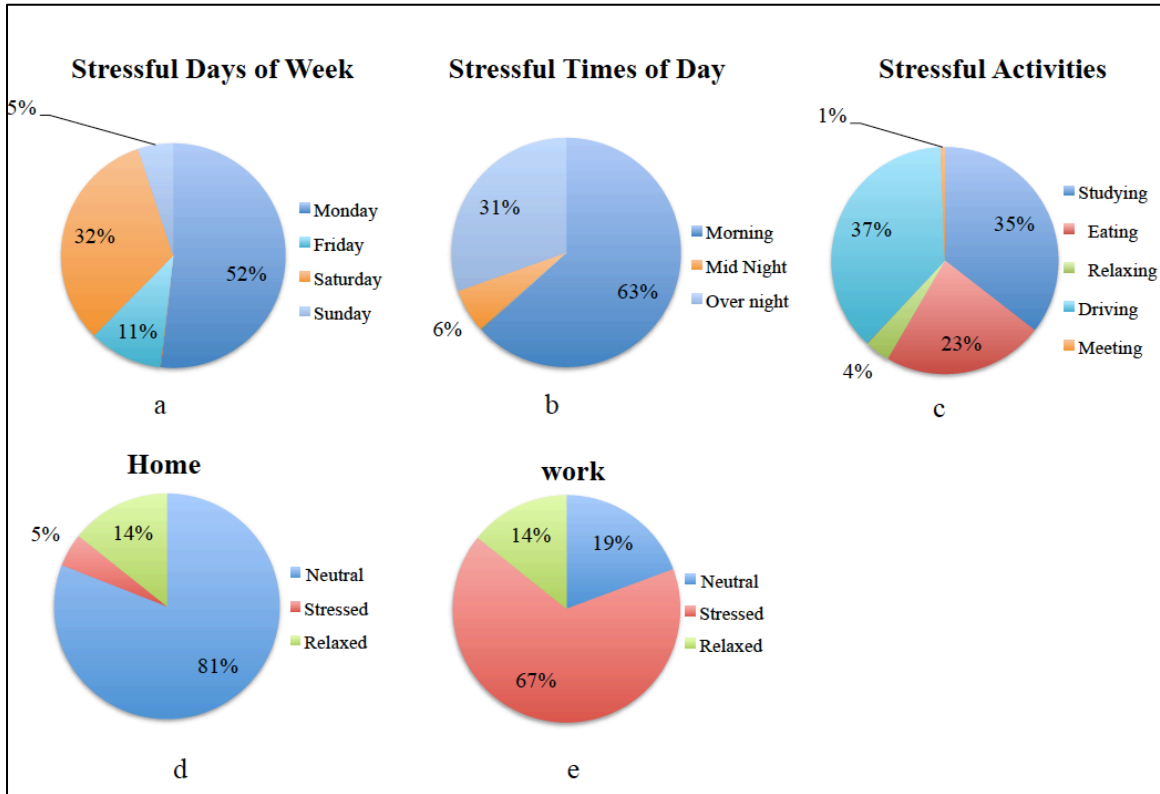


Figure 5.6: Analysis of the Way Contextual Data Affects the Stress Status for User Two. A represents the stressful days of week. B represents the stressful times of the day. C represents the stressful activities. D represents the different stress statuses at home. E represents the different stress statuses at work.

As shown in Figure (5.6), driving and studying are the most stressful activities, Monday is the most stressful day of the week, and morning is the most stressful time of the day. As mentioned in Table (5.1), User 2 was writing a master's thesis during the experimental period, as well as practicing for her driver's license, which explains her high stress status during the experiment.

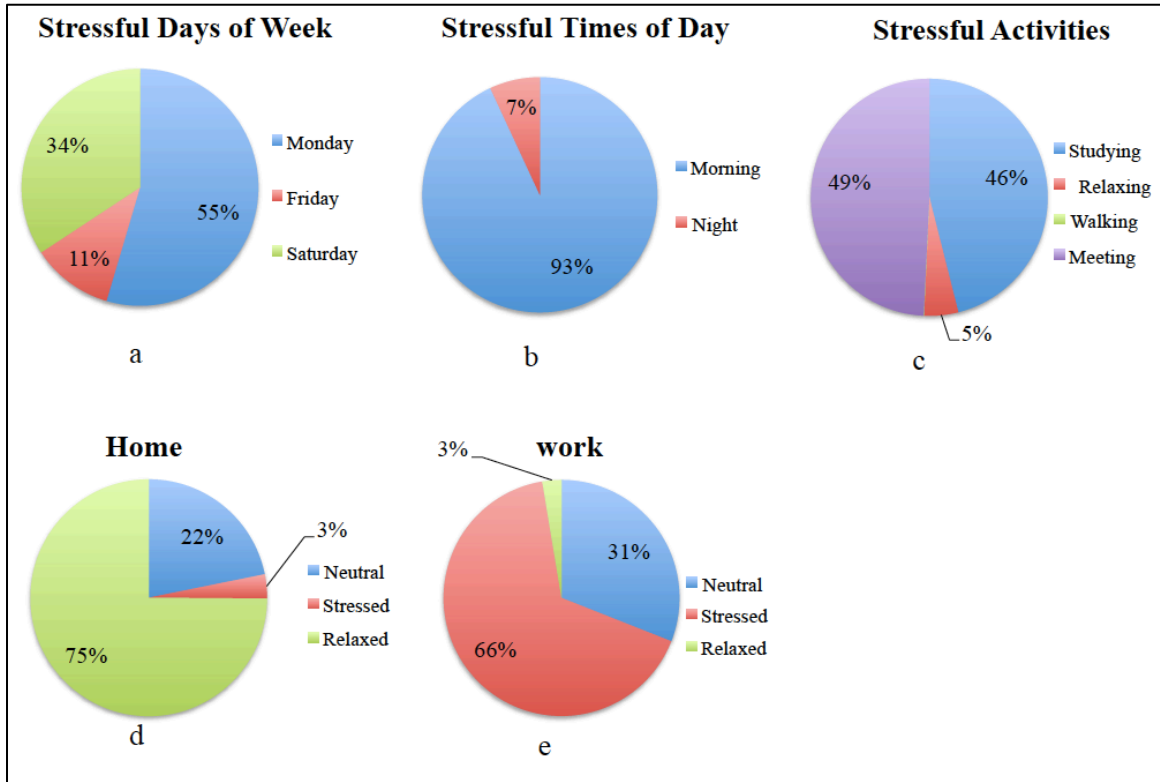


Figure 5.7: Analysis of the Way Contextual Data Affects the Stress Status for User Three. A represents the stressful days of week. B represents the stressful times of the day. C represents the stressful activities. D represents the different stress statuses at home. E represents the different stress statuses at work.

As shown in Figure (5.7), studying and meeting are the most stressful activities, Monday is the most stressful day of the week, and morning is the most stressful time of the day for user three. It also identifies the work location as a stressor.

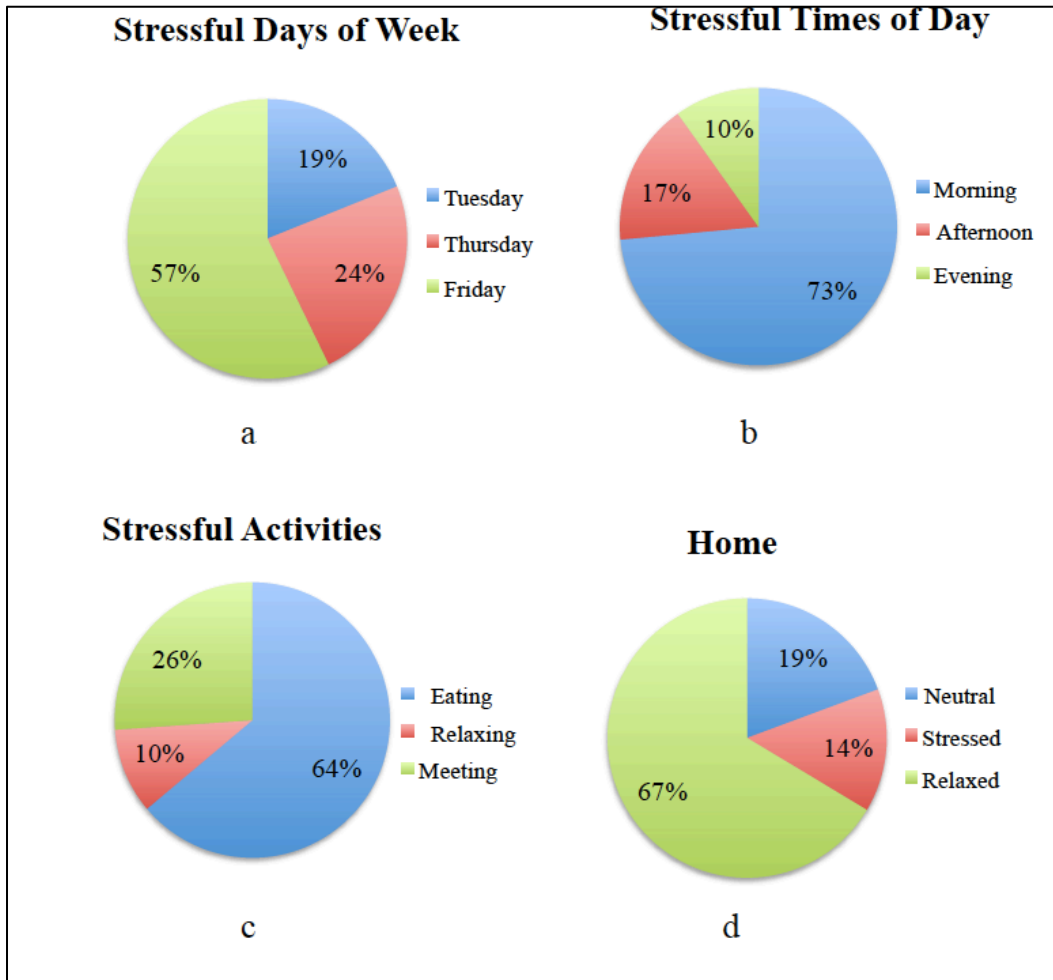


Figure 5.8: Analysis of the Way Contextual Data Affects the Stress Status for User Four. A represents the stressful days of week. B represents the stressful times of the day. C represents the stressful activities. D represents the different stress statuses at home.

As shown in Figure (5.8), eating and meeting are the most stressful activities, Friday is the most stressful day of the week, and morning is the most stressful time of the day for user four.

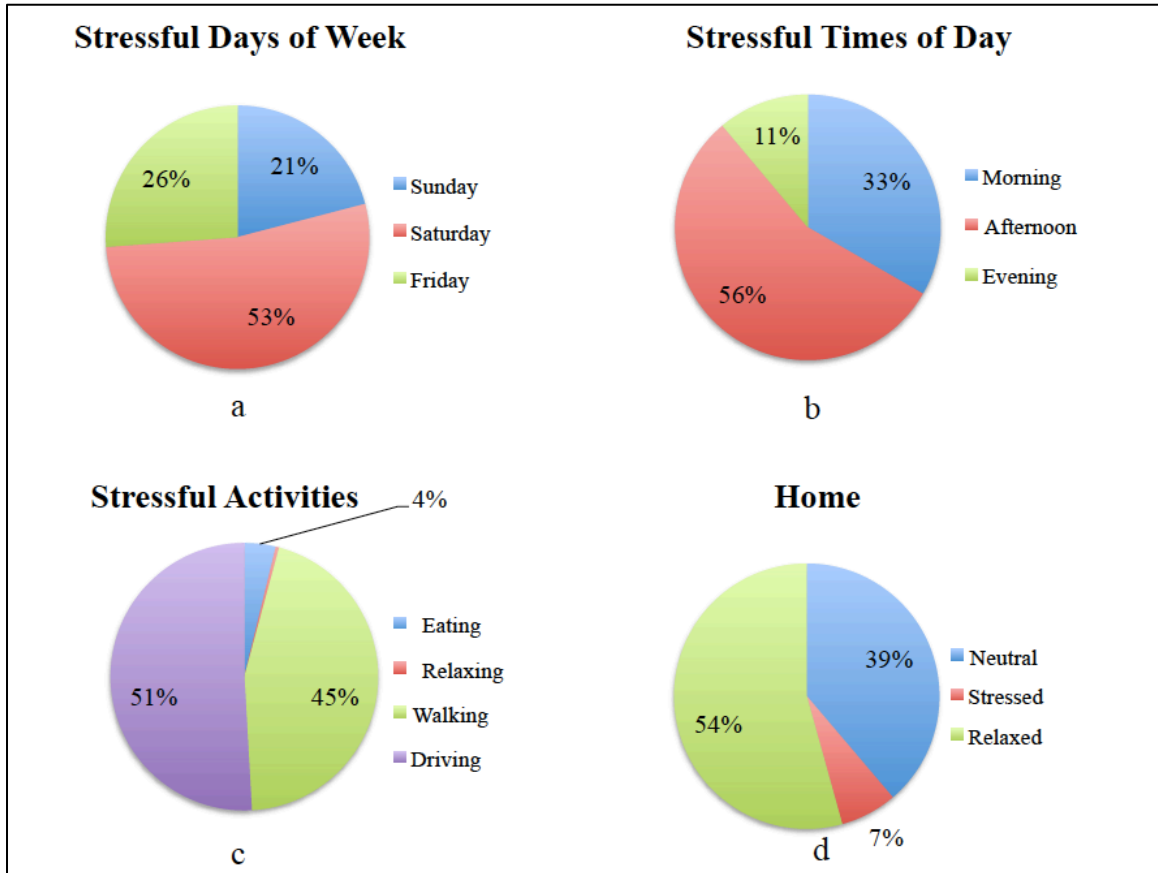


Figure 5.9: Analysis of the Way Contextual Data Affects the Stress Status for User Five. A represents the stressful days of week. B represents the stressful times of the day. C represents the stressful activities. D represents the different stress statuses at home.

As shown in Figure (5.9), driving and walking are the most stressful activities, Saturday is the most stressful day of the week, and afternoon is the most stressful time of the day for user five.

Chapter 6

6 Conclusion and Future Work

Daily stressors we encounter in our environment have a negative impact on our well-being. As most of these surrounding factors are sensible, stress management systems can take advantage of the most frequently used devices, smartphones, as context-aware service providers. In addition, cloud based technology and machine-learning algorithms can also help take stress intervention methods to another level.

Based on the above information, and as explained in this thesis, we propose a context-aware stress prediction system. As stress has a subjective nature and stressors differ from one another, the CASP system is developed to provide a personalized context-aware stress detection algorithm. It recognizes the stress status of the users along with their current context. It incorporates different contextual parameters such as day, time, weather, location, and activity. Moreover, the system predicts the user's stress status based on their current contextual data using the Naive Bayes classifier. The context-aware prediction model adapts to all changes in the environment and in the stress status. Based on the prediction results, relaxation techniques are suggested to the users as a form of early stress intervention, meant to help avoid the stress altogether.

We evaluated the performance of the NB classifier using various accuracy measurements (e.g. TP rate, FP rate, precision, recall, and F-measure).

We conducted a real life non-clinical experiment with five participants. The experimental results demonstrate that the proposed system successfully predicts the user's stress status with an average accuracy of 78.32%. A longer data collection phase will certainly lead to enhanced prediction results. As the collected data increase over time, more stressful situations will be covered.

The independency between the contextual data is one of the proposed system limitations. Another limitation is the number of the contextual data we covered. Another limitation is the number of contextual data we were able to obtain. The method used to collect the information from the users about their activities was not ideal, as the users did not update

their activities very frequently. In future work, the user's calendar could be used to overcome the high-level activity detection limitation.

For future work, we plan to study the relation between different contexts to see how the entire situation can affect a user's stress status compared to the effects of independent contexts. Two or more dependent contexts can play an important role in defining a user's stress status. For example, studying at night might be more stressful than studying in the morning. Moreover, additional features can be added to the system by expanding the contextual data selection, which would provide a richer and more valuable prediction for the users. For example, the system could identify stress makers by connecting the user with nearby people via Bluetooth and marking them as stressors, based on their effect on the user's stress status.

References

- [1] Ahmed, M. U. (2010). A Case-Based Multi-Modal Clinical System for Stress Management. School of Innovation, Design and Engineering, Mälardalen University, Sweden.
- [2] Al Osman, H., Eid, M., and El Saddik, A. (2014). U-Biofeedback: A Multimedia-Based Reference Model for Ubiquitous Biofeedback Systems. *Multimedia Tools and Applications*, 72(3), 3143-3168.
- [3] Al-Aidarroos, K. M., Bakar, A. A., and Othman, Z. (2012). Medical Data Classification with Naive Bayes Approach. *Information Technology Journal*, 11(9), 1166-1174.
- [4] Alam, M. G. R., Cho, E. J., Huh, E. N., and Hong, C. S. (2014). "Cloud Based Mental State Monitoring System for Suicide Risk Reconnaissance Using Wearable Bio-sensors". In *Proceedings of the 8th International Conference on Ubiquitous Information Management and Communication*. 56.
- [5] American Institute of Stress. Effects of Stress. Retrieved from: <http://www.stress.org>.
- [6] Andersson, G., Bergström, J., Holländare, F., Carlbring, P., Kaldo, V., and Ekselius, L. (2005). Internet-Based Self-Help for Depression: Randomised Controlled Trial. *The British Journal of Psychiatry*, 187(5), 456-461.
- [7] Arora, S., and Bhattacharjee, J. (2008). Modulation of Immune Responses in Stress by Yoga. *International journal of yoga*, 1(2), 45-55.
- [8] Bernardi, L., Wdowczyk-Szulc, J., Valenti, C., Castoldi, S., Passino, C., Spadacini, G., and Sleight, P. (2000). Effects of Controlled Breathing, Mental Activity and Mental Stress With or Without Verbalization on Heart Rate Variability. *Journal of the American College of Cardiology*, 35(6), 1462-1469.
- [9] Bonoris, P. E., Greenberg, P. S., Castellanet, M. J., and Ellestad, M. H. (1978). Significance of Changes in R Wave Amplitude During Treadmill Stress Testing: Angiographic Correlation. *The American journal of cardiology*, 41(5), 846-851.
- [10] Brody, D. S. (2003). Improving The Management of Depression in Primary Care. *Disease Management & Health Outcomes*, 11(1), 21-31.
- [11] Burns, M. N., Begale, M., Duffecy, J., Gergle, D., Karr, C. J., Giangrande, E., and Mohr, D. C. (2011). Harnessing Context Sensing to Develop a Mobile Intervention for Depression. *Journal of medical Internet research*, 13(3).
- [12] Carlbring, P., Bohman, S., Brunt, S., Buhman, M., Westling, B. E., Ekselius, L., and Andersson, G. (2006). Remote Treatment of Panic Disorder: A Randomized Trial of Internet-Based Cognitive Behavior Therapy Supplemented with Telephone Calls. *American Journal of Psychiatry*, 163(12), 2119-2125.

- [13] Cinaz, B., Arnrich, B., La Marca, R., and Tröster, G. (2013). Monitoring of Mental Workload Levels During an Everyday Life Office-Work Scenario. *Personal and ubiquitous computing*, 17(2), 229-239.
- [14] Cohen, S., Janicki-Deverts, D., and Miller, G. E. (2007). Psychological Stress and Disease. *Jama*, 298(14), 1685-1687.
- [15] Cohen, S., Kessler, R. C., and Gordon, L. U. (Eds.). (1995). *Measuring Stress: A Guide for Health and Social Scientists*. Oxford University Press.
- [16] Colombo, R., Mazzuero, G., Soffiantino, F., Ardizzoia, M., and Minuco, G. (1989). A Comprehensive PC Solution to Heart Rate Variability Analysis in Mental Stress. In *Computers in Cardiology 1989, Proceedings*. 475-478.
- [17] Diyankova, I. and Asheville, N. Yoga for Emotional Well-Being. (2010).
- [18] Elliott, G. R., and Eisdorfer, C. (Eds.). (1982). *Stress and Human Health: Analysis and Implications of Research: A Study* by The Institute of Medicine National Academy of Sciences. 81(5), New York. Springer
- [19] Evans, G. W. (1984). *Environmental Stress*. CUP Archive.
- [20] Gritti, I., Defendi, S., Mauri, C., Banfi, G., Duca, P., and Roi, G. S. (2013). Heart Rate Variability, Standard of Measurement, Physiological Interpretation and Clinical Use in Mountain Marathon Runners during Sleep and after Acclimatization at 3480 m. *Journal of Behavioral and Brain Science*, 3(01), 26-48.
- [21] Haward, L. R. (1960). The Subjective Meaning of Stress. *British Journal of Medical Psychology*, 33(3), 185-194.
- [22] Hjortskov, N., Rissén, D., Blangsted, A. K., Fallentin, N., Lundberg, U., and Søgaard, K. (2004). The Effect of Mental Stress on Heart Rate Variability and Blood Pressure During Computer Work. *European journal of applied physiology*, 92(1-2), 84-89.
- [23] Hoque, M. E., McDuff, D. J., and Picard, R. W. (2012). Exploring Temporal Patterns in Classifying Frustrated and Delighted Smiles. *Affective Computing, IEEE Transactions on*, 3(3), 323-334.
- [24] Kabachinski, J. (2011). What's The Forecast for Cloud Computing in Healthcare?. *Biomedical Instrumentation & Technology*, 45(2), 146-150.
- [25] Kageyama, T., Nishikido, N., Kobayashi, T., Kurokawa, Y., Kaneko, T., and Kabuto, M. (1998). Self-Reported Sleep Quality, Job Stress, and Daytime Autonomic Activities Assessed in Terms of Short-Term Heart Rate Variability Among Male White-Collar Workers. *Industrial health*, 36(3), 263-272.
- [26] Kocielnik, R., and Sidorova, N. (2015). Personalized Stress Management: Enabling Stress Monitoring with LifelogExplorer. *KI-Künstliche Intelligenz*, 29(2), 115-122.
- [27] Kononenko, I. (1993). Inductive and Bayesian Learning in Medical Diagnosis. *Applied Artificial Intelligence an International Journal*, 7(4), 317-337.
- [28] Kuo, A. M. H. (2011). Opportunities and Challenges of Cloud Computing to Improve Health Care Services. *Journal of medical Internet research*, 13(3).

- [29] MacLean, D., Roseway, A., and Czerwinski, M. (2013). MoodWings: A Wearable Biofeedback Device for Real-Time Stress Intervention. In *Proceedings of the 6th International Conference on Pervasive Technologies Related to Assistive Environments*. 66.
- [30] Malliani, A., Lombardi, F., and Pagani, M. (1994). Power Spectrum Analysis of Heart Rate Variability: A Tool to Explore Neural Regulatory Mechanisms. *British heart journal*, 71(1), 1-2.
- [31] Mäntyjärvi, J., Himberg, J., Kangas, P., Tuomela, U., and Huuskonen, P. (2004). Sensor Signal Data Set for Exploring Context Recognition of Mobile Devices. In *Proc. of 2nd Int. Conf. on Pervasive Computing (PERVASIVE 2004)*.18-23.
- [32] Mäntyjärvi, J., and Seppänen, T. (2003). Adapting Applications in Handheld Devices Using Fuzzy Context Information. *Interacting with Computers*, 15(4), 521-538.
- [33] McEwen, B. S. (2008). Central Effects of Stress Hormones in Health and Disease: Understanding the Protective and Damaging Effects of Stress and Stress Mediators. *European journal of pharmacology*, 583(2), 174-185.
- [34] Merilahti, J., Mattila, E., Plomp, J., Laine, K., and Korhonen, I. (2009). Short-Term Relaxation Responses to A Voice-Guided Mobile Phone Relaxation Application and Self-Guided Relaxation. In *Information Technology and Applications in Biomedicine, 2009. ITAB 2009. 9th International Conference*. 1-5.
- [35] Mohr, D. C., Burns, M. N., Schueller, S. M., Clarke, G., and Klinkman, M. (2013). Behavioral Intervention Technologies: Evidence Review and Recommendations for Future Research in Mental Health. *General hospital psychiatry*, 35(4), 332-338.
- [36] Musumba, G. W., and Nyongesa, H. O. (2013). Context Awareness in Mobile Computing: A Review. *International Journal of Machine Learning and Applications*, 2(1).
- [37] Miller, N. E. (1975). Clinical Applications of Biofeedback: Voluntary Control of Heart Rate, Rhythm, and Blood Pressure. *New horizons in cardiovascular practice*. University Park Press, Baltimore, 239-249.
- [38] Pantelopoulos, A., and Bourbakis, N. (2008). A Survey on Wearable Biosensor Systems for Health Monitoring. In *Engineering in Medicine and Biology Society, 2008. EMBS 2008. 30th Annual International*
- [39] Peternel, K., Pogačnik, M., Tavčar, R., and Kos, A. (2012). A Presence-Based Context-Aware Chronic Stress Recognition System. *Sensors*, 12(11), 15888-15906.
- [40] Phuong, N. H., and Kreinovich, V. (2001). Fuzzy Logic and Its Applications in Medicine. *International journal of medical informatics*, 62(2), 165-173.
- [41] Radespiel-Tröger, M., Rauh, R., Mahlke, C., Gottschalk, T., and Mück-Weymann, M. (2003). Agreement of Two Different Methods for Measurement of Heart Rate Variability. *Clinical Autonomic Research*, 13(2), 99-102.
- [42] Reisman, S. (1997). Measurement of Physiological Stress. In *Bioengineering Conference, 1997, Proceedings of the IEEE 1997 23rd Northeast*. 21-23.

- [43] Satchwell, B. (2015). Mobile Heart Health. *Australasian Biotechnology*, 25(1), 18.
- [44] Setz, C., Arnrich, B., Schumm, J., La Marca, R., Troster, G., and Ehlert, U. (2010). Discriminating Stress from Cognitive Load Using A Wearable EDA Device. *Information Technology in Biomedicine, IEEE Transactions on*, 14(2), 410-417.
- [45] Seyle, H. (1976). Stress without distress. *Hodder & Stoughton Ltd, 1977. ISBN 10: 0340193018*.
- [46] Singh, M., and Queyam, A. B. (2013). A Novel Method of Stress Detection Using Physiological Measurements of Automobile Drivers. *International Journal of Electronics Engineering*, 5(2), 13-20.
- [47] Solomon, R. C., and Stone, L. D. (2002). On “Positive” and “Negative” Emotions. *Journal for the Theory of Social Behaviour*, 32(4), 417-435.
- [48] Strauss, J., Peguero, A. M., and Hirst, G. (2013). Machine Learning Methods for Clinical Forms Analysis in Mental Health. In *MedInfo*. 1024.
- [49] Sun, F. T., Kuo, C., Cheng, H. T., Buthpitiya, S., Collins, P., and Griss, M. (2012). Activity-Aware Mental Stress Detection Using Physiological Sensors. In *Mobile computing, applications, and services*. 211-230.
- [50] Welch, P. D. (1967). The Use of Fast Fourier Transform for The Estimation of Power Spectra: A Method Based on Time Averaging Over Short, Modified Periodograms. *IEEE Transactions on audio and Electroacoustics*, 15(2), 70-73.
- [51] Zephyr Technology, BioHarness BT. Retrieved from: (<http://www.zephyranywhere.com/products/bioharness-3>).
- [52] Zhang, J., Tang, H., Chen, D., and Zhang, Q. (2012). distress: Mobile and Remote Stress Monitoring, Alleviation, and Management Platform. In *Global Communications Conference (GLOBECOM), 2012 IEEE*. 2036-2041.

Appendix A

	place_id	user_id	place_name	latitude	longitude
	Filter	Filter	Filter	Filter	Filter
1	1	2	Home	45.4160035803294	-75.6724052131176
2	2	2	Work	45.3926601913687	-75.7166874781251
3	3	1	Home	45.4160054631171	-75.6710302457213
4	4	1	Work	45.4230727704034	-75.6829808279872
5	5	3	Home	45.4182901797406	-75.6743826717138
6	6	3	Work	45.423191606375	-75.6841164082289
7	7	4	Home	45.4187933321467	-75.6769294291735
8	8	4	Work	45.4234090402577	-75.7001768052578
9	9	5	Home	21.2552576500069	40.4332672432065
10	10	5	Work	21.4331124642959	40.4902004823089

Figure A.1: Table of Places

	tmpresults_id	user_id	LFOverHF
	Filter	Filter	Filter
1	1	2	1.5270726...
2	2	2	1.7483190...
3	3	2	1.4814470...
4	4	2	1.8450856...
5	5	2	1.5333278...
6	6	2	1.4017673...
7	7	2	1.4014452...

Figure A.2: Example of Benchmark Values Table

	stddev_id	user_id	median	stddev
	Filter	Filter	Filter	Filter
1	1	2	1.4017674	0.22844741
2	2	1	0.6953048	0.63628185
3	3	3	0.49513826	0.03345223
4	4	4	1.1279776	0.4255983
5	5	5	2.9871955	1.248332
6	6	6	1.5255947	1.1694095

Figure A.3: Example of Parameters Table

ID	user_id	session_date	timestamp	LFOverHF	median	stddev	stress_level
Filter	Filter	Filter	Filter	Filter	Filter	Filter	Filter
1	2	2015-06-06 19:31:05	1433634036498	1.32606091394881	1.4017674	0.22844741	NORMAL
2	2	2015-06-06 19:31:05	1433634066341	1.26610710843153	1.4017674	0.22844741	NORMAL
3	2	2015-06-06 19:31:05	1433634096514	1.01696653736113	1.4017674	0.22844741	RELAXED
4	2	2015-06-06 19:31:05	1433634126337	1.88913698142735	1.4017674	0.22844741	STRESSED
5	2	2015-06-06 19:31:05	1433634156355	2.46089779706172	1.4017674	0.22844741	STRESSED
6	2	2015-06-06 19:31:05	1433634186335	2.57149812164101	1.4017674	0.22844741	STRESSED
7	2	2015-06-06 19:31:05	1433634216333	3.03790277507016	1.4017674	0.22844741	STRESSED
8	2	2015-06-06 19:31:05	1433634246326	2.47653176153344	1.4017674	0.22844741	STRESSED
9	2	2015-06-06 19:31:05	1433634276322	2.05649774582123	1.4017674	0.22844741	STRESSED
10	2	2015-06-06 19:31:05	1433634306315	2.5891632019601	1.4017674	0.22844741	STRESSED

Figure A.4: Example of Context-Aware Stress Detection Table

location	fuzzy_work	fuzzy_home	fuzzy_weather	fuzzy_day	fuzzy_time	activity
Filter	Filter	Filter	Filter	Filter	Filter	Filter
(45.416492,-75.67191)	very close:0.394	very close:0.0	sunny:0.005	saturday:0.567	midnight:0.78	Eating
(45.416473,-75.671936)	very close:0.394	very close:0.0	sunny:0.005	saturday:0.566	midnight:0.777	Eating
(45.41668,-75.671745)	very close:0.399	very close:0.0	sunny:0.005	saturday:0.566	midnight:0.777	Eating
(45.416595,-75.671844)	very close:0.397	very close:0.0	sunny:0.005	saturday:0.565	midnight:0.774	Eating
(45.41654,-75.671814)	very close:0.396	very close:0.0	sunny:0.005	saturday:0.565	midnight:0.774	Eating
(45.41655,-75.67185)	very close:0.396	very close:0.0	sunny:0.005	saturday:0.564	midnight:0.771	Eating
(45.41652,-75.67193)	very close:0.395	very close:0.0	sunny:0.005	saturday:0.564	midnight:0.771	Eating
(45.416576,-75.671776)	very close:0.397	very close:0.0	sunny:0.005	saturday:0.563	midnight:0.767	Eating
(45.41653,-75.67193)	very close:0.395	very close:0.0	sunny:0.005	saturday:0.563	midnight:0.767	Eating
(45.416664,-75.67183)	very close:0.399	very close:0.0	sunny:0.005	saturday:0.562	midnight:0.764	Eating

Figure A.5: Example of Context-Aware Stress Detection Table-2

```

1 <data>
2 <s ID="2998" user_id="2" session_date="42177.0883449074" timestamp="1434960884097" stddev="0.22844741" stress_level="STRESSED"
3 <s ID="2999" user_id="2" session_date="42177.130775463" timestamp="1434960890080" stddev="0.22844741" stress_level="RELAXED" l
4 <s ID="3000" user_id="2" session_date="42177.130775463" timestamp="1434960950057" stddev="0.22844741" stress_level="RELAXED" l
5 <s ID="3001" user_id="2" session_date="42177.0883449074" timestamp="1434961034092" stddev="0.22844741" stress_level="STRESSED"
6 <s ID="3002" user_id="2" session_date="42177.130775463" timestamp="1434961070058" stddev="0.22844741" stress_level="RELAXED" l
7 <s ID="3003" user_id="2" session_date="42177.130775463" timestamp="1434961100081" stddev="0.22844741" stress_level="RELAXED" l
8 <s ID="3004" user_id="2" session_date="42177.0883449074" timestamp="1434961154070" stddev="0.22844741" stress_level="RELAXED"
9 <s ID="3005" user_id="2" session_date="42177.0883449074" timestamp="1434961214107" stddev="0.22844741" stress_level="STRESSED"
10 <s ID="3006" user_id="2" session_date="42177.130775463" timestamp="1434961220086" stddev="0.22844741" stress_level="RELAXED" l
11 <s ID="3007" user_id="2" session_date="42177.0883449074" timestamp="1434961244142" stddev="0.22844741" stress_level="STRESSED"
12 <s ID="3008" user_id="2" session_date="42177.130775463" timestamp="1434961250085" stddev="0.22844741" stress_level="STRESSED"
13 <s ID="3009" user_id="2" session_date="42177.0883449074" timestamp="1434961274117" stddev="0.22844741" stress_level="STRESSED"
14 <s ID="3010" user_id="2" session_date="42177.130775463" timestamp="1434961280098" stddev="0.22844741" stress_level="STRESSED"
15 <s ID="3011" user_id="2" session_date="42177.0097685185" timestamp="1434961296108" stddev="0.22844741" stress_level="RELAXED"
16 <s ID="3012" user_id="2" session_date="42177.0883449074" timestamp="1434961304132" stddev="0.22844741" stress_level="STRESSED"
17 <s ID="3013" user_id="2" session_date="42177.130775463" timestamp="1434961310086" stddev="0.22844741" stress_level="STRESSED"
18 <s ID="3014" user_id="2" session_date="42177.0097685185" timestamp="1434961326115" stddev="0.22844741" stress_level="RELAXED"
19 <s ID="3015" user_id="2" session_date="42177.0883449074" timestamp="1434961334129" stddev="0.22844741" stress_level="STRESSED"
20 <s ID="3016" user_id="2" session_date="42177.130775463" timestamp="1434961340009" stddev="0.22844741" stress_level="STRESSED"
21 <s ID="3017" user_id="2" session_date="42177.0097685185" timestamp="1434961356143" stddev="0.22844741" stress_level="RELAXED"

```

Figure A.6: XML Prediction Request

	id numeric	StressLevel text	PredictedStressLevel text	LocationLatitude real	LocationLongitude real	WorkText text	WorkValue real	HomeText text	HomeValue real	WeatherText text	WeatherValue real	DayText text	DayValue real
1	863	NORMAL	NORMAL	45.4167	0.67168	very close	0	very close	0	sunny	0.005	friday	0.377
2	6	STRESSED	STRESSED	45.4165	0.67185	very close	0.396	very close	0	sunny	0.005	saturday	0.564
3	7	STRESSED	STRESSED	45.4165	0.67193	very close	0.395	very close	0	sunny	0.005	saturday	0.564
4	8	STRESSED	STRESSED	45.4166	0.671776	very close	0.397	very close	0	sunny	0.005	saturday	0.563
5	9	STRESSED	STRESSED	45.4165	0.67193	very close	0.395	very close	0	sunny	0.005	saturday	0.563
6	10	STRESSED	NORMAL	45.4167	0.67183	very close	0.399	very close	0	sunny	0.005	saturday	0.562
7	11	STRESSED	STRESSED	45.4166	0.671814	very close	0.397	very close	0	sunny	0.005	saturday	0.562
8	12	STRESSED	STRESSED	45.4165	0.67185	very close	0.396	very close	0	sunny	0.005	saturday	0.561
9	13	STRESSED	STRESSED	45.4165	0.6719	very close	0.394	very close	0	sunny	0.005	saturday	0.561
10	14	STRESSED	STRESSED	45.4166	0.67182	very close	0.397	very close	0	sunny	0.005	saturday	0.56
11	15	STRESSED	STRESSED	45.4165	0.67188	very close	0.395	very close	0	sunny	0.005	saturday	0.56
12	16	STRESSED	STRESSED	45.4164	0.672104	very close	0.393	very close	0	sunny	0.005	saturday	0.559
13	17	STRESSED	STRESSED	45.4166	0.671814	very close	0.396	very close	0	sunny	0.005	saturday	0.559
14	18	STRESSED	STRESSED	45.4165	0.67191	very close	0.395	very close	0	sunny	0.005	saturday	0.558
15	19	NORMAL	STRESSED	45.4165	0.6719	very close	0.396	very close	0	sunny	0.005	saturday	0.558

Figure A.7: Example of Context-Aware Prediction Results Table