*C02029: Doctor of Philosophy*

*CRICOS Code: 00099F*

*33847 PhD Thesis: Software Engineering*

*June 2022*

# *AutoMAP*

## *Smart Wearables for Depression: Toward Autonomous Mental Health Monitoring for the Elderly*

*Fiza Tariq Mughal*

School of Computer Science

Faculty of Engineering & IT

University of Technology Sydney

NSW - 2007, Australia

# AutoMAP
# Smart Wearables for Depression: Toward Autonomous Mental Health Monitoring for the Elderly

*A thesis submitted in fulfilment of the requirements for the degree of*

Doctor of Philosophy
*in*
Software Engineering

*by*

**Fiza Tariq Mughal**

*to*

School of Computer Science
Faculty of Engineering and Information Technology
University of Technology Sydney
NSW - 2007, Australia

June 2022

# Abstract

Depression has increasingly become a significant concern over the years, even more so since the start of the COVID-19 pandemic. In the first year of the pandemic, anxiety and depression increased by 25% in all populations. Many older-aged people believe that mental illness is not something that can be treated and is often stigmatized. More so, older-aged people often have an increased risk of other mental issues, misdiagnosis as general age-related deterioration such as general cognitive impairment, Alzheimer's disease, or dementia. This increased dependence on others and loss of general socio-economic status increases the likelihood of depression. In older aged people, mental illness and other age-related conditions inevitably lead to caregiver burden and dependence. One approach to aid in potentially alleviating caregiver burden while allowing older-aged people to live alone, is the use of consumer grade smart wearables to monitor depressive tendencies.

Increasingly, studies are being conducted on emotion recognition using contemporary technologies, particularly smart wearables. There is however a lack of focus on older-aged people (65+ years) as the target demographic. When this cohort has been investigated, study materials consist of obtrusive on-body sensors or are administered in lab-based settings with external emotion stimulation. As such, these are inapplicable or less accurate in real-world settings.

This thesis aims to address the above issues by proposing AutoMAP - Autonomous

Mental health monitoring for older Aged Persons - a framework that incorporates consumer grade smart wearables for implicit user input and machine learning to infer depressive tendencies in order to create this system, a framework was proposed based on a comprehensive literature review in the area of emotion recognition and machine learning. This literature review was conducted to develop a clearer overview of mental health issues around the world, particularly in older-aged people, and the current state of research in emotion recognition and smart technologies. Through an in-depth analysis of existing research, strengths and limitations of existing approaches were highlighted. An assessment of these strengths and limitations allowed for the refinement of the current research project, aiming to reduce some of the identified limitations. Second, a conceptual framework that provides autonomous mental health monitoring for older aged people was developed. Third, an experimental study was performed with mixed-design methods to validate the proposed framework. Lastly, data outputs from the experimental study were analysed to validate the feasibility of the framework. Activity data from wearable devices were used to train and evaluate machine learning models to assess their predictive performance toward depressive tendency detection.

Based on the above, a proof-of-concept platform was created that consisted of (1) implicit user input, (2) machine learning, and (3) emotion reporting infrastructure. Various machine learning models will be explored for predictive modelling on device data. This research aims to conceptualize, propose, and pilot the impact and potential applicability of AutoMAP. These findings can collectively build a foundation toward better mental health monitoring and independent living for older aged users.

# Author's declaration

I, *Fiza Tariq Mughal* declare that this thesis, submitted in fulfilment of the requirements for the award of Doctor of Philosophy, in *Software Engineering*, *Faculty of Engineering and IT* at the University of Technology Sydney, Australia, is wholly my own work unless otherwise referenced or acknowledged. In addition, I certify that all information sources and literature used are indicated in the thesis. This document has not been submitted for qualifications at any other academic institution.

SIGNATURE: ______________________________

[Fiza Tariq Mughal]

DATE: $30^{th}$ June, 2022

PLACE: Sydney, Australia

# Dedication

*To my grandmother and mother for inspiring this project . . .*

# Acknowledgments

My first and foremost expression of gratitude is towards Allah for blessing me with strength, perseverance and patience throughout life, especially in the past 3 years.

I would like to thank my primary supervisor, Dr. Jaime Garcia, for the constant support and encouragement. He has been a confidant and a mentor that I have had the immense honor to learn from - a true maestro. During my Ph.D., Dr. Jaime Garcia has continuously challenged me to keep at it, while also enforcing mental health breaks when I myself didn't think I needed them. I owe much of my success to him.

I would like to thank my co-supervisors Dr. William Raffe and Dr. Peter Stubbs for the countless hours of mentoring and proofreading my publications and thesis. Their comments and suggestions on my project, publications and thesis were integral to my success.

A special thanks to Professor Ian Kneebone, for providing invaluable feedback in the design process for this project. And to Professor Valerie Gay for providing integral feedback on my research since my very first assessment.

I want to thank Irina and Sharmila for proofreading this thesis and being a huge part of my support system. To Rabeea, Kiara, Nissa, Hamza, Mauro, Carlos, Heather, and Phoebe for being the most supportive friends one could ask for and for keeping me sane especially during the final stages of my PhD.

A big thanks to the participants, especially those that helped recruit more participants for my study. Without their contributions, this project would not have gotten far.

I'd also like to thank my supervisors for my workspace in the Games Studio. This place became my second home for 3 years, and with it came another little support system (I'm looking at you Ezwan, Paddy, Andres and Elena).

Last but most definitely not least, an exceptional thanks to my family. To my parents for all the sacrifices they made to allow me to fulfill this dream, and for lovingly encouraging me to pursue paths that I feel passionate about. To my mother for being one of my biggest role models. To my sister and brother-in-law for all the pampering they've indulged me in and being my comfort space, and to my cousins for being the OG entourage. Most importantly, to my grandparents, without whom the idea for this project would not have existed.

2.48. IYKYK

# List of Publications

## Journals

1. **F. Mughal**, W. Raffe, P. Stubbs, I. Kneebone, J. Garcia, *Fitbits for Monitoring Depressive Symptoms in Older Aged Persons: Qualitative Feasibility Study*, **JMIR Publications (accepted)**.

2. A. O. Kingsley, U. G. Inyang, O. Msugh, **F. Mughal** and A. Usoro, *Recognizing facial emotions for educational learning settings*, **IAES** International Journal of Robotics and Automation **(accepted)**.

## Conferences

3. **F. Mughal**, W. Raffe and J. Garcia, *Emotion Recognition Techniques for Geriatric Users: A Snapshot*, 2020 IEEE 8th International Conference on Serious Games and Applications for Health - **SeGAH** '20 **(accepted)**.

4. **F. Mughal**, W. Raffe, P. Stubbs and J. Garcia, *Towards depression monitoring and prevention in older populations using smart wearables: Quantitative Findings*, 2022 IEEE 10th International Conference on Serious Games and Applications for Health - **SeGAH** '22 **(accepted)**.

## Others

5. **F. T. Mughal**, *Latest Trends in Human Activity Recognition and Behavioral Analysis using Different Types of Sensors*, 345th International Conference on Innovative Engineering Technologies - **ICIET (accepted)**.

# Table of Contents

# List of Figures

**Figure** **Page**

# List of Tables

# List of Abbreviations

| | |
|---|---|
| AD | Alzheimer's Disease |
| API | Application Programming Interface |
| AutoMAP | Autonomous Mental Health Monitoring for Older Aged Persons |
| CSV | Comma Separated Value |
| ECG | Electrocardiogram |
| EDA | Electrodermal Activity |
| EEG | Electroencephalography |
| EMA | Ecological Momentary Assessment |
| EMG | Electromyography |
| ER | Emotion Recognition |
| ETL | Extract-Transform-Load |
| FTD | Frontotemporal Dementia |
| GSR | Galvanic Skin Response |
| GDS | Geriatric Depression Scale |

| | |
|---|---|
| HC | Healthy Controls |
| HRV | Heart Rate Variations |
| MDD | Major Depressive Disorders |
| MAE | Mean Absolute Error |
| MAPE | Mean Absolute Percentage Error |
| MSE | Mean Square Error |
| MCI | Mild Cognitive Impairment |
| MVPA | Moderate-to-vigorous Activity Minutes |
| NIA | National Institute on Ageing |
| R2 | R-square |
| RF | Radio Frequency |
| REM | Random Eye Movement |
| RSP | Respiration |
| RMSD | Root Mean Square Deviation |
| RMSLE | Root Mean Squared Logarithmic Error |
| SKT | Skin Temperature |
| SR | Speech Recognition |
| VR | Voice Recognition |
| WHO | World Health Organization |

# Chapter 1

# Introduction

This chapter provides an overview of the problem that has led to this research. Also presented, are the aims, research questions, scope, and rationale of the project. The structure of this thesis is also detailed at the end of this chapter.

## 1.1 Depression, emotions and perception

Older-aged people experience stressful situations similarly to younger people, as well as stressors encountered later in life such as a continuing loss in capabilities and declining set of functional abilities. Additional stressors include a reduced socioeconomic status due to retirement. These experiences can lead to psychological distress or deeper mental health issues for older aged people. In addition, older aged people are also more vulnerable to various sorts of psychological, verbal, or physical abuse, which will increase the likelihood of anxiety or depression. This makes it difficult or impossible for family members to be at ease when their older relatives are living alone. Furthermore, some older-aged people insist on staying on their own regardless of functional (dis)ability.

Despite this, there are also older-aged people that have no preference in their living arrangements, though their families may be unable to assist due various reasons, such as loss of income or being faced with additional expenses that may not necessarily be covered under insurance (Grunberg, 2010). Patients with cognitive impairment-type diseases such as Alzheimer's or Dementia face deterioration of mental functions as well as potentially changing personality traits. This leads to caregiver burden as carers may not know how to assist the person (Hamdy et al., 2017) or know the best management of their relatives. There is an increasing number of older aged people living alone (Hsu & Chien, 2009). To tackle this issue, enhanced monitoring systems that assist older aged people to live alone with minimal to no caregiver burden, could be used. A potential approach could be the implementation of a more implicit and independent monitoring system that uses Emotion Recognition (ER) methods.

ER is the processing and perception or identification of human emotion. This is often performed by facial cues or explicit verbal expression. While this is something most humans can do, there are ongoing studies that attempt to automate this process (Ko, 2018; Xiaoxi et al., 2017). Some diagnoses where humans tend to find difficulty perceiving or expressing emotions include Autism, Alzheimer's, Parkinson's, dementia, or general age related cognitive deterioration. In the case where the use of or the ability of verbal expression is not hindered, older-aged people are still known to report depression more infrequently than younger people, which could lead to increased anxiety (Sas et al., 2017). The aim of the current study is to propose an emotion recognition-based monitoring system that requires minimal explicit user input or interaction, to potentially aid and assist older aged people that live alone by reducing mental or caregiver burden. Such a system could allow older-aged people to live alone, subsequently increasing their independence but also allow for timely assistance in the event of an emergent situation. It would also reduce caregiver burden by lowering continual human monitoring and reliance/dependence. There are various studies on how on-body or wearable devices can be used to assist users to a better quality of life through positive physical activity behavioural changes. Many studies focus on using 'on-body' sensors or wearable devices such as smartwatches or smartphone sensors in ongoing attempts towards smarter emotion recognition (Kumar et al., 2021; Luxton et al., n.d.; Park et al., 2020; Shu et al., 2020). However, these studies often require an explicit form of user input, such as being in a set frame of reference for facial recognition or requiring on-body sensors such as electroencephalography (EEG) or electrocardiography (ECG) devices. While younger users adapt to emerging technologies and devices, older-aged people tend to have difficulties implementing new technologies, leading to apprehension, frustration, and avoidance of newer technologies. Some research (Kasteren et al., 2017; Vaportzis et al., 2017) also contends that modern technologies, like smart homes, are met with apprehension from the older aged people due to a lack of familiarity and comfort. This is further strengthened when the user experiences cognitive deterioration with conditions such as Alzheimer's or Dementia (Koo & Vizer, 2019).

Growing focus has been placed on determining the connection between explicit physiological signals and implicit internal feelings (Kasteren et al., 2017; Khundaqji et al., 2020; Quiroz et al., 2018; Rohani et al., 2018; Shu et al., 2020). There are many studies on emotion recognition using various kinds of sensors (Ali et al., 2018; Patlar Akbulut et al., 2020). Studies highlight that motion based or wearable approaches are not necessarily the best fit for monitoring as some of these studies attach various sensors to the body. Other studies have pre-set classification models which only apply to younger subjects and would likely be ineffective for older aged people. This is because older aged people, particularly those with cognitive impairments or chronic illnesses tend to be more unpredictable in their behaviours (Grunberg, 2010; Virtanen et al., 2017), which would be classified as outliers in current classification models.

The ability to infer emotions from facial expressions becomes more difficult with age related cognitive impairment (Virtanen et al., 2017). For speech recognition, we must consider the possibility that older aged people become non-verbal due to age related conditions such as a stroke. Furthermore, recognition of positive emotions is less impaired with a severe decline in cognitive function as opposed to negative emotions, i.e., test participants had a higher accuracy when inferring positive emotions such as happiness (69.3%) in comparison to anger (50.8%) and sadness (45.8%). Similar research (Wiechetek Ostos et al., 2011) also observed that there is selective impairment in 'disgust' and 'fear' recognition with increasing cognitive impairment due to progressive damage to neural structures, linked to emotion and facial recognition. It is plausible that as people become older it becomes harder for a person to infer and express their own emotions. For example, a study (Park et al., 2017) showed Facial Emotion Recognition impairment in patients with FrontoTemporal Dementia (FTD), Alzheimer's Disease (AD), and Mild Cognitive Impairment (MCI). To do this, the study examined the Facial Emotion Recognition (FER) performance of patients with FTD, AD, and those with MCI against healthy controls (HCs). Recognition of negative emotions differentiated between participants with FTD and those with MCI, AD or HCs. The recognition of positive emotions showed no differences. They also stated that there is still a need for more enhanced emotion recognition tools. There have been studies on exploiting smart home technologies for activity recognition. However, only one study focused on the older aged people (Kasteren et al., 2017). The study focused on activity recognition only and required explicit user input and interaction to some extent. The study commonly found false triggers in the algorithm due to an inability to handle exceptions caused by variations in the daily routines of older aged people.

## 1.2 Implications and Limitations of existing research

Although there is a lot of research being conducted on the applications of smart wearables for health and emotion recognition, not enough focus has been placed on the older-aged people as the target demographic. Current studies on emotion recognition using smart wearables are commonly conducted in lab-based settings with obtrusive on-body sensors and require emotion elicitation through external stimulation. This may render the outputs of these investigation less applicable in real-world implementation. Existing research has most commonly produced binary, categorical outputs such as happy-neutral-sad or happy vs sad. In this thesis, we introduced the AutoMAP framework for Autonomous mental health Monitoring of older-aged people. This framework consists of three components that come together to produce an implementation that would provide less obtrusive mental health monitoring focused on the older-aged people, with data collected in real-world routine-based settings. Figure 1.1 illustrates a comparison of current emotion recognition research versus autonomous mental health monitoring (AutoMAP), which is introduced and detailed further in this thesis.

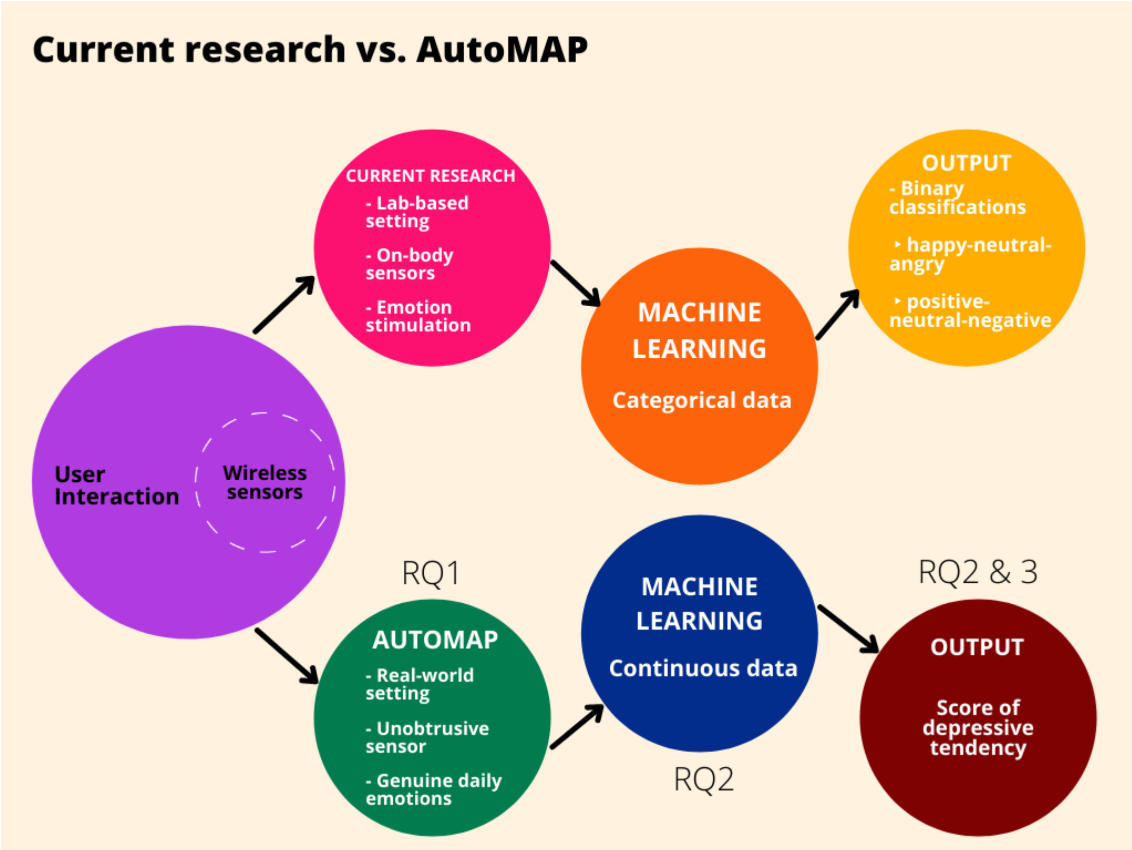



Figure 1.1: Proposed autonomous monitoring vs current state of research

## 1.3 Research Stakeholders

In the context of this research, stakeholders refer to the combined group of people that can benefit from the findings of this project. There are two subgroups of stakeholders; Older aged people using the wearable activity monitor, and their caregivers that will be the emotion reporting mobile application end users. Users of the wearable serve as stakeholders in that their mental health would be monitored without necessarily needing to explicit interactions with caregivers or medical professionals. Smartwatch users consist of 173 million users in 2022 alone (Fig. 1.2), with numbers increasing to 253 million by 2025 (Laricchia, 2019). Subsequently, aged care communities, assisted ambient living societies, and healthcare providers with outpatient care may potentially benefit from this research.

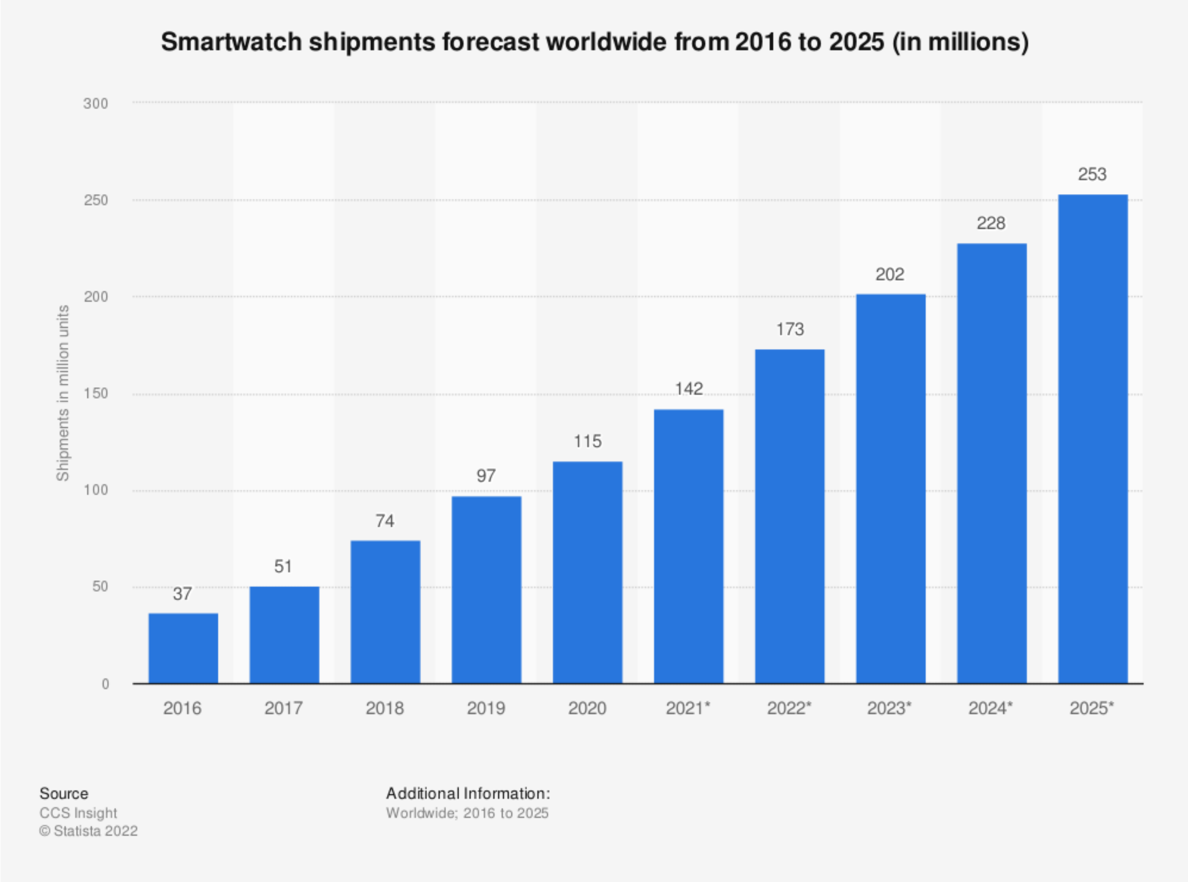



Figure 1.2: Smartwatch shipments forecast worldwide from 2016 to 2025 (in millions) (Laricchia, 2019)

## 1.4 Research Aims & Objectives

### 1.4.1 Primary Aim

Current research incorporating smart wearables has a positive effect on physical activity and behavioural change (Chum et al., 2017; Lyons et al., 2017; Tong et al., 2018), while emotion recognition through smart wearables is effective in binary classifications across happy-neutral-sad / positive-neutral-negative (Quiroz et al., 2018; Shu et al., 2020; Zhang et al., 2016). However, many of these studies are performed in lab-based settings, require emotion stimulation, or for the participant to wear obtrusive on-body sensors (Fig. 1).

Table 1.1: Research Overview.

| Research Question | Research Aim | Research Objectives |
|---|---|---|
| RQ1. Will older aged people use the chosen wearable device and find it easy and convenient in a routine-based implementation? | Evaluate feasibility and convenience of a least obtrusive data collection device. | Conduct comprehensive literature review.<br><br>Develop conceptual framework for autonomous mental health monitoring.<br><br>Determine the best suited device for unobtrusive data collection.<br><br>Carry out data collection.<br><br>Gather feedback on feasibility and convenience of unobtrusive data collection method. |
| RQ2. Can physiological data from wearable devices be indicative of depressive tendencies in users? | Train a classifier to process physiological inputs into a score outlining depressive tendency in the user. | Perform ETL processes to prepare dataset for model training.<br><br>Evaluate model performance. |
| RQ3. Can the proposed framework have a potential impact in reducing caregiver burden and assisting older aged people to live independently? | Evaluate potential framework impact pre- and post- implementation. | Conduct post-study interviews with returning participants for feedback on feasibility and acceptability of framework. |

The primary goal of this research project is to (1) explore the relationship between wearable device data and depression and (2) the applicability of machine learning models to infer depressive states in users. Consequently, this research will introduce an unobtrusive autonomous mental health monitoring framework for older-aged people. This could reduce undiagnosed or reported mental health incidents in older-aged people. Additionally, this may also reduce caregiver burden and dependence. This framework aims to predict depressive tendencies in users and is not intended as a diagnostic tool. For this, there are three research questions, which will be addressed through three aims and corresponding objectives. These aims focus on developing and validating a proposed conceptual framework to monitor mental health among older aged people. It is intended that the thesis objectives will benefit the research stakeholders.

### 1.4.2 Objectives

The first part of this research will focus on determining the feasibility of using consumer grade smart wearables as an unobtrusive method to collect data from the user, while requiring minimal explicit interaction with the device. Comparing existing studies, the sensors, approaches they applied, and limitations encountered, will help determine which device is the most convenient for the user. The second part of this research aims to develop a machine learning model that uses physiological data from smart wearables. The model would process device data and produce a predictive score of depressive tendency in the user. Extracted physiological device data and user questionnaire responses will be used for preparing the dataset to train and test the machine learning model. The third and final part of this research intends to validate the feasibility and acceptability of using consumer grade smart wearables in predicting a user's depressive state, while potentially alleviating caregiver burden and dependence that sometimes occurs as a secondary consequence of aging.

To evaluate the feasibility and convenience of using a consumer grade smart wearable to collect physiological user data, the first objective is to conduct a comprehensive literature review to provide insight on what methods, sensors, and devices are being used in recent studies and we identify any gaps in the literature. These objectives together, will (1) aid in selecting a device for the study that is minimally obtrusive, convenient, cost-effective, and easily implementable and (2) lead to the development of a conceptual framework for autonomous mental health monitoring for older aged people.

The second set of objectives consist of extract-transform-load (ETL) processes to form a data pre-processing pipeline that leads to a novel dataset consisting of physiological Fitbit data and questionnaire responses. This dataset will be used to train, test, and validate machine-learning models to determine the most suitable model for the conceptual framework.

The final objective will validate the feasibility and potential impact of the framework for older aged people and their caregivers. This framework should assist in monitoring the depressive tendencies of a user and automatically report to a caregiver should depressive tendencies reach a concerning threshold. A summary of the research questions, aims and objectives is presented in Table 1. Figure 1.1 provides an overview of current research in relation to the aims and objectives of this research.

## 1.5 Research Significance

The aim of this research project is to evaluate the applicability of wearable devices and machine learning to aid and assist older aged people that live alone (Fig. 1.3), especially those with depressive tendencies, cognitive impairments, or general age-related cognitive decline. In some cases, older aged people are non-verbal or are unable to express or perceive emotions. Some older aged people are also less likely to vocalize mental health issues like depression, and it is often misdiagnosed as general cognitive decline (CDC, 2021). This research focuses on detecting and monitoring tendencies in older-aged people through unobtrusive means and minimal explicit user input or interaction. This study will be using smartwatches to collect the user's walking, sleep, and heart rate data to derive their emotional state (depression centred). Figure 1.3 illustrates the proposed mental health monitoring framework - AutoMAP. This framework will be discussed in further detail in Chapter 2.

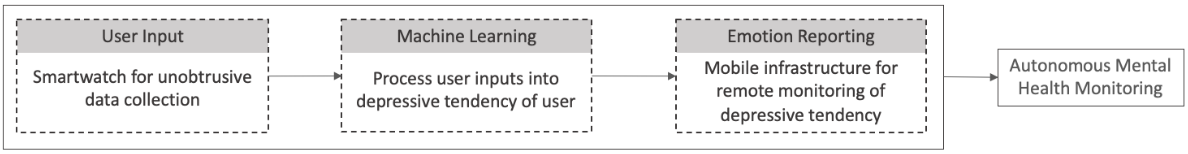



Figure 1.3: AutoMAP framework

This research intends to improve the quality of life and reduce caregiver burden by assisting the older aged end users of the proposed framework in being more independent, while being monitored and possibly having emergent care available when necessary. Validation was be performed through pre- and post-trial interviews with users, older aged people, along with evaluation of predictive performance of the ML model.

## 1.6 Approach

The methodology of this research project follows a mixed qualitative and quantitative design.

Prior to study commencement, literature reviews, device selection, framework conception along with ethical considerations were addressed and ethics approvals were acquired. Following this, participants were recruited for the study for a 4-week period. As this research was performed in human subjects, there is always some level of potential risk, however all foreseeable risks were addressed and accounted for, with mitigation techniques if required. During the 4-week period, data were collected through the selected device, along with a set of online questionnaires. Participants were instructed on the frequency of questionnaire response and the requirements for keeping the device on though the study period.

Finally, upon the completion of the 4-week study period and post-completion interviews, qualitative and quantitative analyses were be conducted. The outcomes of these analyses helped determine the feasibility of the framework, select the most suitable machine learning model, and effectively build a foundation toward autonomous mental health monitoring by means of minimal user interaction or obtrusion. A summary of study phases and corresponding outcomes are shown in Table 2.

## 1.7 Structure of this Dissertation

The remainder of thesis is structured as follows: Chapter 2 presents a comprehensive literature review into emotion recognition, depression, current trends in research on smart wearables and mental health, the gaps, and limitations of existing literature.

Chapter 3 details the methods and protocols applied in this study. This chapter explains the processes and considerations taken that led to the final study design, the types of analyses performed along with the rationale for choosing each type of analysis. Chapter 4 presents the qualitative findings from interviews conducted to evaluate the feasibility, acceptability, and usability of the data collection device. Chapter 5 provides a descriptive overview of the extract physiological smart wearable data, along with an in-depth look at the quantitative findings of the study. This chapter describes the types of machine learning models evaluated in the model selection stage, along with a deeper dive into the best performing model. Challenges and limitations faced with the dataset and quantitative analysis as a whole are also presented.

Finally, Chapter 6 concludes this thesis, explaining how the combined study outcomes validate the conceptual framework and build a foundation toward a proof-concept

Table 1.2: Research Phases and Aimed Outcomes.

| Phase | Outcomes |
|---|---|
| Phase 1: Pre-intervention considerations | Literature review.<br><br>Device selection.<br><br>Framework conception.<br><br>Carry out data collection.<br><br>Ethical considerations and approvals. |
| Phase 2: Participant recruitment and onboarding | Dissemination of study materials.<br><br>Conduction of pre-intervention interviews. |
| Phase 3: Feasibility of using consumer grade wearables for data collection with older aged users | Data collection using wearable device.<br><br>Conduction of post-intervention interviews.<br><br>Assessment of framework feasibility. |
| Phase 4: Data extraction and analysis | Measures of agreement and feasibility of framework.<br><br>Development of extract-load-transform pipeline for Fitbit data. |
| Phase 5: Phase 5: Development of predictive model | Pre-processing of combined and individual participant datasets.<br><br>Machine learning model selection and evaluation. |
| Phase 6: Validate proof-of-concept framework | Evaluation of applicability and applicability of the conceptual framework.<br><br>Analysis of results and conclusions. |

framework. This chapter also presents concluding statements, limitations, and promising future directions of this research.

## 1.8 Conclusion

This chapter introduced the research questions, aims and objectives while setting the foundation of this research. This chapter presented an overview and discussion on the current state of emotion recognition research. The next chapter sets out the theoretical foundations of this work and presents the current state of research in emotion recognition and smart technologies.

# Chapter 2

## Literature Review.

This chapter provides an overview of current emotion recognition techniques and why they may not necessarily be suitable or feasible for older-aged people. There is a lot of research on implementing smart devices and wireless sensors for emotion recognition in users. While this is an important direction toward mental health monitoring, not enough focus is placed on older aged people as the focal cohort. This may present as an obstacle for real world implementation, as these experiments are often held in lab-based settings and consist of younger study participants. With age, there is an increased likelihood of experiencing cognitive decline, which may cause impairments in daily functional capabilities (Murman, 2015). One of the main cognitive functions affected by age is attention, along with perception among other sensory capacities (Glisky, 2007).

This analysis serves as a foundation for a proposed conceptual framework toward an autonomous monitoring system for older aged people which could minimize the need for explicit user input or interaction while still monitoring well-being. A stronger understanding of existing literature and a comparison of the strengths and limitations will supplement the decision-making process in deciding the type of smart wearable, algorithms and accompanying technologies that will be incorporated within the AutoMAP framework introduced in Chapter 1.

## 2.1 Background

### 2.1.1 Emotion Recognition

Emotion recognition (ER) is the processing and perception or identification of human emotion. This is commonly performed either by facial cues or by explicit verbal expression. While this is something most humans do innately, there are ongoing studies attempting to automate this. Sometimes humans find it difficult perceiving or expressing emotions, for example if they have Autism, or age induced cognitive deterioration such as Alzheimer's, Parkinson's Disease or dementia. In cases where the use of or the ability of verbal expression is not hindered, older-aged people report depressed tendencies less than younger people do, leading to increased anxiety (Depression and Older Adults, n.d.). Currently, there are various studies focusing on using on-body sensors or wearable devices such as smartwatches or smartphone sensors in an ongoing attempt toward smarter emotion recognition (Y. Chen & Shen, 2017; Ghosh & Riccardi, 2014; Lustrek et al., 2015; Quiroz et al., 2018). However, these studies require explicit forms of user input, such as being in a set frame of reference for facial recognition or requiring an on-body sensor such as an Electroencephalography (EEG) or Electrocardiography (ECG) device. While younger users adapt to emerging technologies and devices quickly, the older aged people tend to face challenges in this respect, leading to apprehension or even frustration and avoidance of using these technologies as a whole (Yaddaden et al., 2016). Some focus has been placed on determining the connection between explicit physiological signals and implicit internal feelings such as fear (Ali et al., 2018; Miranda et al., 2021;

Shu et al., 2018). Rohani et al. (2018) conducted a review of 46 studies evaluating the correlation between objective behavioural features and depressive mood symptoms in participants. The review showed significant correlations between behavioural aspects and depressive moods, with albeit inconsistent study methodologies. There are studies on emotion recognition using various kinds of sensors for physiological signals. The ability to infer emotions from facial expressions becomes more difficult with age related cognitive impairment, and at a higher rate in the case of deteriorating cognitive function (Virtanen et al., 2017). This may also lead to apprehension or even frustration and avoidance of new technologies in general (Boustani et al., 2007). The study also found that recognition of positive emotions is less impaired with a severe decline in cognitive function as opposed to negative emotions. Test participants had a higher accuracy when inferring positive emotions such as happiness (69.3%) in comparison to anger (50.8%) and sadness (45.8%). Wiechetek Ostos et al. (2011) additionally observed that there is selective impairment in disgust and fear recognition with increasing cognitive impairment due to progressive damage to neural structures linked to emotion and facial recognition. Park et al. (2017) performed tests and found further evidence on Facial Emotion Recognition (FER) impairment in patients with FrontoTemporal Dementia (FTD), Alzheimer's disease (AD), and those with mild cognitive impairment (MCI). To do this, the study examined the FER performance of patients with FTD, AD, and those with MCI against healthy controls (HCs). Their approach found that recognition of negative emotions differentiated between participants with FTD and those with MCI, AD or HCs. The recognition of positive emotions showed no differentiation. They also stated that there is still a need for more enhanced emotion recognition tools.

There have also been studies on exploiting smart home technologies for activity recognition. Kasteren et al. (2017) focused on older aged people however, the study focused on activity recognition only, and required a level of explicit user input and interaction. There were also cases of false triggers due to an inability to handle exceptions caused by variations in the daily routines of the older aged people. Shu et al. (2018) state that while physiological data is a more definitive method for determining a person's emotional state, there is still a certain level of difficulty for inferring a person's emotional state using a single physiological signal. It was hypothesized that combinations of various physiological sensors could lead to improvements in emotion recognition approaches. They outperformed a comprehensive review on various current physiological signal-based emotion recognition approaches. The review evaluated EEG, ECG, HR, GSR, RSP, and EMG methods proposed by researchers for emotion recognition. Smoothing filters and additional noise extraction methods were also used to remove any interference or background noise such as respiration sinus arrhythmias (RSA) from RSP or eye blinks from EEG signals. The review assessed emotion recognition frameworks as well as common setups for high quality physiological data acquisition. Also considered were the conditions under which study participants were required under, for valid data collection, such as sitting motionless in front of a screen for visual emotion elicitation. The review also reported challenges encountered by researchers using physiological signals for emotion recognition, starting off with emotion elicitation. They found that the process of emotion elicitation is commonly undertaken in a lab-based setup, with the test subject sitting motionless in front of a screen with emotion triggering stimuli being played for them.

The remainder of this chapter looks into the various types of sensors used in current emotion recognition studies. Insight into ageing, technology acceptance and mental health is presented as well. While under the same umbrella as wearable devices, on-body devices refer to relatively obtrusive devices such as EEG headsets and chest-mounted monitors. On the other hand, smart devices refer to wearables that are deemed less obtrusive. Wireless devices are included under non-contact devices. The majority of

Table 2.1: Summarized grouped information on studies based on the type of device.

| Device Type | Author | Year | Multi-modal | Measures |
|---|---|---|---|---|
| On-body | Wan et al. | 2021 | Y | ECG, EDA, SKT |
| | Tengfei Song et al. | 2019 | N | ECG |
| | Egger | 2019 | Y | ECG, EMG |
| | Shu et al. | 2018 | Y | ECG, EMG, GSR, RSP |
| | Benaissa et al. | 2017 | Y | Motion, environment |
| | Soroush et al. | 2017 | N | EEG |
| | Kone et al. | 2017 | Y | ECG, EMG, GSR |
| | Jamil et al. | 2015 | N | Gait |
| Smart | Buda et al. | 2021 | Y | Step count, EMA |
| | Miranda et al. | 2021 | Y | ECG, GSR, SKT |
| | Shu et al | 2020 | N | HR |
| | Jaiswal | 2018 | Y | Gait, sleep |
| | Lee et al. | 2018 | N | Gait, HR |
| | Zhang et al. | 2016 | Y | Gait |
| Wireless | Adib | 2019 | N | RF |
| | Barrett et al. | 2019 | N | Facial expression |
| | Rozanska et al. | 2018 | Y | Verbal-visual context |
| | Zhao, Adib & Katabi | 2016 | N | RF |
| | Kadir et al. | 2014 | N | EMF |
| | Schneider et al. | 2014 | N | Gait |
| | Wang et al. | 2014 | N | Facial expression |
| | Weichetek Ostos et al. | 2011 | N | Facial expression |

the studies adopted non-contact measurements such as facial recognition, followed by on-body contact devices such as chest mounted heart rate monitors (Fig. 2.3a). Other studies focused on smart devices such as smart homes (noncontact) and smart watches (contact). Below is a further in-depth analysis of the selected papers. Figure 2.1 shows an overview of how the two main categories are further split up in this section.

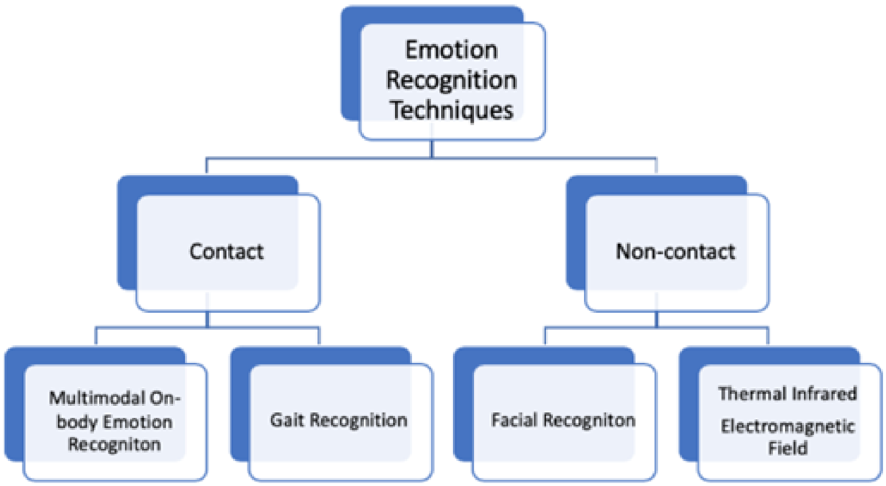



Figure 2.1: Categorization of Emotion Recognition Techniques

### 2.1.2 Mental Health and Caregiver Burden

Previously, studies have found that the older-aged people are less like to address mental health issues (CDC, 2021; Hasin  Link, 1988). Older-aged people are less likely to consider depression as a mental disorder, and as such are less likely to recount depressive episodes (Hasin & Link, 1988). Hasin et al. (1988) conducted a survey to assess whether major depression was recognized as mental disorder among older-aged people. Their study found that respondents did not consider major depression to be a psychological or emotion challenge. Where a diagnosis is sought, it is not uncommon for older-aged people to get mental health ailments such depression misdiagnosed as a general age-related change in emotions (CDC, 2021). This is particularly concerning as older-aged people experience stressful encounters that are also encountered by the younger generation, as well as stressors encountered later in life such as loss in capabilities and a declining set of functional abilities (WHO, 2021). Inactivity has also shown to result in increased depression (Zimmermann et al., 2019). Additional stressors could even include socioeconomic status reductions due to retirement. These experiences can lead to psychological distress or deeper mental health issues for the older aged people (WHO, 2021). In addition, older-aged people are also more vulnerable than younger cohorts to various sorts psychological, verbal, or physical abuse among others (WHO, 2021); which could also result in anxiety or depression. With age there is also a growing likelihood of being riddled with cognitive impairment diseases, such as Alzheimer's or Dementia, face deterioration of mental functions as well as a likelihood of changing personality traits.

This makes it difficult for family members to be at ease when older relatives and friends live alone. Some older-aged people insist on living alone as they would like to be independent. Despite this, there are also some older-aged people that have no preference on their living situation, though their families may be unable to take care of them (Grunberg, 2010). This potentially leads to caregiver burden as family members try to determine how to best care for or how to assist the person (Hamdy et al., 2017). Caregiver burden can be categorized as a multidimensional series of negative responses that may occur when serving as a primary carer (Liu et al., 2020). This could potentially increase the likelihood of depressive disorders in older-aged people.

There is generally an increasing number of older-aged people living alone worldwide (Hsu & Chien, 2009) with the global population statistics register showing a reduced birth rate and an increasing life expectancy (De Nardin et al., 2020). Introducing an enhanced monitoring system that allows the older-aged people to live alone with minimal to no caregiver burden or dependence may possibly assist in bettering the quality of life of older aged people. A potential approach could be the implementation of a more implicit and independent monitoring systems that use Emotion Recognition (ER) methods. With this manuscript we aim to analyse and synthesize the most common emotion recognition techniques being used by researchers. The analysis will focus on common methods, as well as the feasibility and convenience of use for the study subjects, ambient settings, and whether the study uses a multimodal ER approach. Most importantly we look at the applicability of the respective methods with older aged people as main study subjects.

### 2.1.3 Ageing and Technology

There is a lot of research using smart devices for health applications (C. Chen et al., 2021; Domin et al., 2021; Hernandez-Cruz et al., 2019; J. Naslund et al., 2014). Studies have shown that the older-aged people are happy and accepting of technologies that are easy to setup and generally unobtrusive (Sas et al., 2017). This provides a sense of autonomy and independence. Older-aged people are welcoming of technologies that will assist them in staying physically and mentally active, and prefer technologies that are adaptive, smart, and not designed exclusively for older aged people (Voit et al., 2016). In Australia, technologies that have the potential to prolong independence such as Smart Wearable Systems (SWS) are highly accepted (Chan et al., 2012). These systems have the capacity to gather a range of sensory data that can be used to monitor the environments surrounding older-aged. These systems could potentially assist in various patient treatments, however challenges with signal processing, data analysis and interpretation remain.

Steering toward age-related cognitive impairment, Koo & Vizer (2019) conducted a systematic review on the use of mobile technology for people with dementia. Their review provides evidence that mobile technologies can be used to partially compensate for the reduced functionality as a result of dementia. This further highlights the potential impact of smart devices for mental health monitoring in older aged people. Looking into the factors surrounding positive user experiences in Assisted Ambient Living (AAL), a qualitative study found that older-aged people and their informal support networks would be more receptive to technologies that are implemented within their daily lives, especially if they would allow them to maintain their current living situation (Elers et al., 2018). To this, the remaining section looks discuss different smart technologies being used in current research for emotion recognition. A discussion leading up to a selected device for the AutoMAP framework will also be presented.

## 2.2 Article Selection

### 2.2.1 Databases Searched

Relevant literature was found using databases containing peer-reviewed articles in the fields of health, computer science, information technology, affective computing, and mental health. Databases that were included were PubMed, JMIR, ACM Digital Library, IEEE Explorer, Science Direct.

### 2.2.2 Search Criteria

Results were filtered through certain criteria to be selected for further analysis. Search terms included the main ideas; older-aged people, geriatric, and emotion recognition. To further narrow down on relevant papers extra keywords were added such as cognitive impairment. Some keywords were added for search exclusion purposes, such as education and speech. This was done as a large amount of search results (n = 7,352) included education and speech evolution centred literature. Two different searches were performed, one for general emotion recognition studies and one on emotion recognition with geriatrics. The main search was first performed in 2020 with the following two search strings, with an updated search in 2022. After the initial search, papers were manually filtered to find those most relevant to this review.

Initial search string:

emotion recognition AND !"education"AND !"speech".

Older cohort and depression centred search string:

"emotion recognition" OR "mental health" OR "depression" AND "smart *" AND "elder*" OR "geriatric" AND "cognitive impairment" AND !"education" AND !"speech".

Results of both strings were further filtered to publication dates within the past 10 years to capture the current state of the research. The most relevant papers were on emotion recognition techniques such as facial recognition, gait recognition, and physiological sensors. The AND keyword in the search string entails necessary keyword combinations and can be replaced with parentheses where applicable or necessary, as shown below. The OR keyword entails a keyword or set of keywords separate from those grouped with the AND keyword, that may also match the search requirements.

Alternative structure:

("emotion recognition", "mental health", "depression"), "smart *", ("elder*", "geriatric", "cognitive impairment"), (!"education", !"speech")

### 2.2.3 Information Analysis

From the initial search, articles were manually filtered by reading through the abstract and conclusion to find papers focusing on emotion recognition techniques, the conditions under which the studies were performed (lab-based or real-time setting), method of user data collection (contact versus non-contact) as well as the potential applicability to older aged people. Duplicates were removed. We then performed an in-depth analysis of remaining papers by removing articles that were not as relevant as initially assumed. We then looked into the contributions, practices, limitations and challenges of the remaining papers.

Ninety-five relevant articles were found. These articles will be further evaluated and discussed, highlighting the strengths and limitations and how they relate to the research questions introduced in Chapter 1. The next section provides a detailed look into current relevant emotion recognition studies and the methods and approaches they adopted. The literature is grouped based on similarities in objectives and methods used.

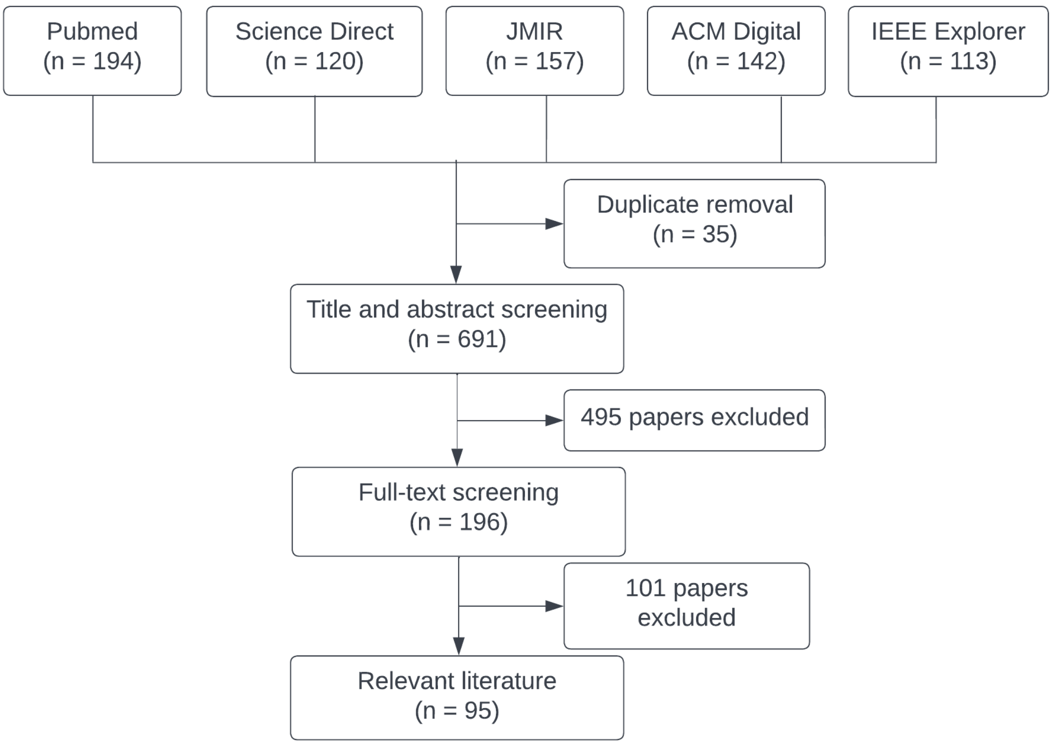



Figure 2.2: Article selection (post-preliminary filtering)

## 2.3 Emotion Recognition Works

### 2.3.1 Contact Emotion Recognition

#### 2.3.1.1 Multimodal Emotion Recognition

Zangeneh Souroush et al. (2017) focuses more on the use of various physiological methods for emotion recognition, such as online versus offline recognition, on-body measurements (EEG, ECG) as well as emotion stimulation approaches detailed in the methods for studies reviewed in their study. The review summarises recent studies on emotion recognition with a greater focus on EEG based studies. It also reiterated that there was no consensus on the nature of emotions. Additionally, environmental variations led to physiological changes that will likely affect emotions. Along the physiological sensor route, Kone et al. (2017) detected emotions using musical therapy with three physiological sensors for a multimodal emotion recognition approach. They used music of various genres to trigger or alter different emotions alongside Galvanic Skin Response (GSR), Electromyography (EMG), and ECG. The multimodal approach grouped the emotions into three categories: neutral, joy, and pleasure; with the highest accuracy being 93% for the Pleasure classification. The study showed an improved recognition rate using a multimodal approach in comparison to other current emotion recognition approaches. While the classification method demonstrated accuracy in emotion recognition, this method still requires on-body devices to collect data and infer emotions.

Benaissa et al. (2018) states that heart rate and breathing were used in various studies for emotion recognition as they are strong indicators of emotion. While other emotion recognition approaches are inefficient in monitoring older-aged people, this method is more applicable for this population. The paper notes that facial thermal imaging is an ongoing challenge for emotion recognition; their own proposed approach makes use of heart rate, breathing and thermography together for a more efficient multimodal approach focusing emotion recognition in older-aged people. In their method, Benaissa et al. (2018) proposed positioning the sensor on the subject's chest for activity recognition. For emotion recognition the optimal sensor placement was on both sides of the human chest to collect breathing data using the heart rate. They also propose using thermography for facial thermal imaging. Their proposal suggests using a combination of sensors for enhanced emotion recognition. The method can be imagined working as a comfortable wearable instead of an on-body chest strap on device.

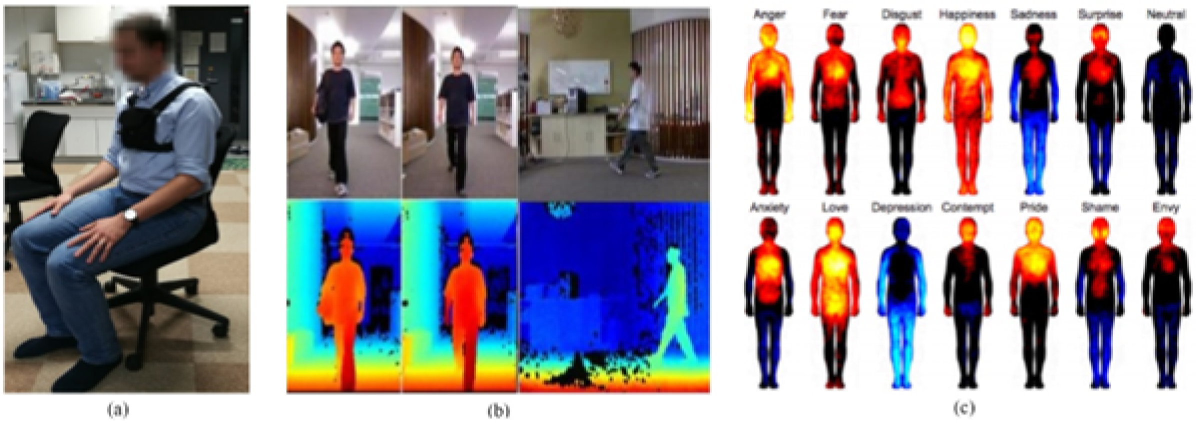



Figure 2.3: (a) On-body chest mounted sensor (b) In-house gait recognition (c) Body Atlas of heat distribution with changing emotions

#### 2.3.1.2 Gait Recognition

While there is a lot of facial emotion recognition methods and studies, researchers have observed that recent studies in have shown a relationship in a person's emotions and the way they walk (Quiroz et al., 2018; Zhang et al., 2021). Quiroz et al. (2018) collected gait and heart rate data to propose an enhanced emotion recognition approach, outside of a potentially mood-altering laboratory setup. The aim of the study was to infer participant emotions using smartwatch sensor data. To analyse the relationship between a person's emotional state and their gait, participants walked with a smartwatch as well as a heart rate monitor. Participants watched audio-visual clips and listened to audio (happy/sad) stimuli to alter their emotions and corresponding gait. The study provided evidence that a person's emotions can be inferred using movement sensor data. However, the approach required further validation in order to definitively state that movement sensor data can be used for emotion recognition.

In an earlier study, Jamil et al. (2015) attempted to use gait analysis to detect emotions in Autistic children. Autistic patients, like Alzheimer's or dementia patients are often not the focus of emotion recognition studies. The study required children to wear bodysuits with markers on them, which was intolerable for some of the autistic children. The children moved erratically, out of the assigned camera frame and removed the markers from their bodysuits. While Autistic children are not the focus of the study at hand, it relates in the target group (the older-aged people) possibly being non-verbal or unable to determine and express their own emotions. As mentioned earlier, Grunberg (2010) stated that children and some older aged people share characteristics in their lack of independent capabilities. Both age groups require similar care and a relatively similar level of dependence on caregiver. Some older-aged people can also be unpredictable with their movements and emotional responses to stimuli (Hamdy et al., 2017). The study by Jamil et al. (2015) added that there are earlier studies that considered gait recognition. However, none of these have focused on children with autism. Similarly, there does not appear to be any studies using gait-based emotion recognition with older-aged people as test participants, to date.

### 2.3.2 Contactless Emotion Recognition

Body language serves as an integral element of nonverbal communication which is normally perceived before expressions (Schneider et al., 2014). More studies are focusing on wireless methods for data collection, leading to emotion classification in an unobtrusive manner. This is increasingly important, especially with older aged people (Sas et al.,

2017) who are most satisfied with unobtrusive devices and least satisfied with bulky or complicated devices (Sas et al., 2017). One of the most common approaches in the literature discussed so far is facial recognition. There is, however, a growing interest toward using non-contact methods to collect emotion recognition data, for example, thermal imaging or electromagnetic radio waves (Adib, 2019). As mentioned earlier, Adib (2019) used wireless Radio Frequencies (RF) to detect human activity through walls, as well as infer user emotions. Fig. 2.3b shows sample images of how non-contact gait patterns are tracked. Further discussed are studies focusing on non-contact emotion recognition.

#### 2.3.2.1 Facial Recognition

As a potential improvement, Rozanska et al. (2018) proposed an embedded system that implemented various emotion recognition methods in an Internet of Things (IoT) device for remote emotion detection. Their proposed system consisted of a robot device with a computer vision camera. When a person or multiple people approach the robot, it would detect and infer their emotions. As the system is meant to be an IoT setup, it also uses sound, video, and speech-to-text analysis and recognition, including the choice of words and tone. They used body language and mimic expressions to infer emotions and found that some positive emotions, such as happiness, are detectable at a longer distance. Negative emotions, such as sadness, require a shorter distance for accurate detection. Anger was better detected through body posture as opposed to facial mimicry features. While the system is accurate, it requires between 14 to 20 seconds for emotion recognition and classification.

#### 2.3.2.2 Thermal Infrared/Electromagnetic Field

Some researchers are focusing on studying thermal infrared images for emotion recognition. Leber (2014) mapped heat distributions in the human body based on people's emotions. Based on heat distributions, anger is felt more in a person's head, while positive emotions like happiness or love are spread throughout the body (Nummenmaa et al., 2014). Negative emotions like depression and sadness on the show a deactivation of sensations in comparison to other less negative emotions. The article was based on an earlier study by Finnish researchers on the 'Body Atlas', which demonstrates how different emotions are manifest in the body. Fig. 2.3c shows the aforementioned 'Body Atlas'. The study proposes the use of the Bolztmann machine for emotion recognition using thermal infrared facial images. Their method outperformed other approaches using temperature statistic features or handcrafted features. Kadir et al. (2014) noted that the human electromagnetic field (EMF) changes with varying activities and health. Their study used EMF readings to distinguish left hemisphere stroke patients and found that left hemisphere stroke patients have lower frequencies on the left side in comparison to the right side. Kadir et al. (2014) also showed that both left and right hemisphere stroke patients have significantly lower electromagnetic radiations (EMR) in comparison to non-stroke patients. Similarly, Ahmad et al. (Ahmad et al., 2020) collected electromagnetic radiation and aura analysis to determine physical and psychological fitness in people with and without Down syndrome. Zhao et al. (2018) exploited the positive aspects of audio-visual and physiological sensor-based approaches and proposed a new method to achieve viable outcomes in emotion recognition, without the limitations of on body or audio-visual sensors. To do this, they use Radio Frequency (RF) signal reflections off the human body to infer their body movements and emotions. Their new algorithm extracted heartbeat data from wireless signals which were then provided to a machine learning emotion classifier, leading to emotion recognition. The study also notes that while smartphones are also used for emotion classification; they require a lengthy time frame to personalize the analyses. The algorithm proposed in their study operates on minute-scale intervals for emotion recognition (Zhao et al., 2018). This study is seemingly the first to exploit RF signals from body reflections and infer user emotions. It is one to follow closely as the objective is close to, if not almost the same as that of the approach being introduced in the current study.

## 2.4 Discussion

While other studies have pre-set classification models that may apply for younger subjects, they may be inefficient for older-aged people. This is because older-aged people, especially those with cognitive impairments tend to be more unpredictable in their behaviors than younger people (Hamdy et al., 2017), which could result in outliers for current trained classification models. These "unpredictable" behaviors however can be anticipated, avoided and defused (Hamdy, 2017). Ideally, classfication models trained on this data should aim to take these movements into account and provide persoanlized outputs as opposed to a 'one-size-fits-all' approach. It could be plausible that with age, it becomes more difficult to express their own emotions. Hence, researchers need to further analyse physiological changes to study and classify emotions, while keeping accounting for the likely artefacts that could obstruct data collection. Many of the studies discussed above have introduced new approaches, multimodal approaches or enhanced versions of previous approaches for emotion recognition. Having said that, there are still a few issues that have yet to be considered. One such issue is that older-aged people are often not tech friendly as opposed to the younger people that can adapt much quicker (Vaportzis et al., 2017).

### 2.4.1 Contact Emotion Recognition

While there are currently various emotion recognition studies, they come with certain limitations. A physiological sensor such as an ECG monitor can inadvertently alter a person's mood by hindering everyday activities and becoming frustrating as a body attachment. Quiroz et al. (2018) also similarly stated that emotions are not necessarily a conscious function and are a reaction to surrounding stimuli or ambience. Laboratory based setup ups with large on-body sensors such as EEG devices or chest straps, can serve as obstructive devices inadvertently leading to unwanted mood alterations. Further, lab-based setups rely heavily on the stimuli provided to trigger or provoke specific emotions (Shu et al., 2020). Research (Drew, 2018; Nature, 2019) has also highlighted that while emotions are shown in various ways such as voice, body language, gait or face, strict laboratory-based setups may obstruct the validity of the collected data. Shu et al. (2020) reiterated the requirement for more work toward creating more natural, realistic ambience for data collection in order to create accurate genuine emotions.

The second issue was subjective responses to external stimuli and general emotion production. Different people have different emotional responses to different scenarios and stimuli, due to which there is yet to be a fully tried and tested approach for emotion state inference through physiological signals. Additionally, there were different physiological responses with varying intensities of emotion stimulation (Schneider et al., 2014). This serves as an obstacle along the way of individualised, real-time applications of physiological methods for emotion recognition. An additional issue is the small population sizes for test groups each study, resulting in small data sets. Hence, there is a need for larger scale studies for physiological sensor-based emotion recognition.

### 2.4.2 Non-contact Emotion Recognition

People express anger, disgust, fear, happiness, sadness, and surprise differently across cultures, situations, and even between people in a single situation. More so, a set classification for a scowl may not detect what a person is actually feeling, it is plausible that it may communicate something other than the detected emotions. Considering that facial movements demonstrate many emotions, there is a need for more studies that focus on examining how facial movements actually vary alongside other social inputs that lead to a conclusive foundation on how humans perceive emotions (Wang et al., 2014). In the case of voice or speech recognition, we must consider the increased likelihood that older-aged people may be non-verbal due to age related conditions such as stroke. While facial recognition can gather data to infer emotions wirelessly, it requires the person to be within a set camera frame. Older aged people, especially those with cognitive impairment such as Alzheimer's or Dementia, are often unpredictable and atypical in their behaviors (Hamdy et al., 2017), which could be another obstacle using a facial recognition approach. This also includes explicit user interaction with the emotion recognition method or system. Aside from requiring explicit user input or interaction, facial recognition depends entirely on outward expression. Expressions contribute approximately 55% to the effect of the message being conveyed, with vocal tone contributing 38% and vocal cues contributing 7%.

Based on this, they performed experiments focusing on facial emotion recognition using images. The method improved the AAL experience for the older-aged people. However, the proposed method still requires further enhancement for to increase accuracy as well as take video inputs instead of images. Some researchers (Rozanska et al., 2018) stated that while facial recognition is of particular importance in non-verbal communication, it is a mistake to focus solely on facial emotion recognition. Shu et al. (2020) stated that physical signals such as facial expressions or speech are not heavily reliable as it is fairly easy for people to control their outward physical expressions. This can be demonstrated when a person smiles while being in a negative emotional state. However, facial thermal imaging (Benaissa et al., 2018; Wang et al., 2014) is still being used frequently in current emotion recognition studies. Temperature statistical are adopted and features extracted from facial regions of interest. Using visual inputs, another study demonstrated that using gait recognition for human identification (Fig. 2.3b) and created an in-house gait recognition database (Christie, n.d.). While their study does not focus on emotion recognition using gait, it shows the potential of using gait recognition through non-contact means.

The majority of emotion recognition studies focus of facial emotion recognition in younger people with only a few focusing on the older-aged people (Benaissa et al., 2018; Kasteren et al., 2017; Park et al., 2017; Yaddaden et al., 2016). These studies discussed that while smart homes are intended to aid older-aged people with their daily needs, the technologies involved are unfamiliar to older-aged people, who are often uncomfortable with the implementation of these devices. In order to bridge this gap, robots are proposed to act as the intermediary between the sensors/devices and the older-aged people, though introducing robots may result in a larger technology barrier. For this to be effective, the robot needs to be at the same level of communication and understanding as the user. In an attempt to improve this method of the AAL smart home experience, the study a used facial emotion recognition and found that facial expressions are an unreliable mode of emotion recognition. One reason for this is a phenomenon called 'facial mimicry' shows that when watching movie clips, or given a stimulus to alter emotions, expression of the same emotion will appear on the persons face. While this shows empathy, it does not facilitate emotion recognition in people in the absence of external stimuli. Similarly, outward expressions can be falsified easily or not reflect what a person is actually feeling (Barrett et al., 2019; Shu et al., 2020). Studies show that facial recognition is still the most studied emotion recognition approach among the literature discussed so far. Other approaches, as mentioned earlier, rely on audio-visual or physiological data collected through cameras, on body sensors or wearable devices.

There is an clear gap in implicit emotion recognition, as well as emotion recognition approaches being implemented or tested with the older-aged people as the target population.

## 2.5 The AutoMAP Conceptual Framework

Based on the literature review, we decided to investigate wearable devices such as smartwatches, as these are relatively unobtrusive compared to other common data collection methods. Physiological inputs such as gait patterns, sleep patterns and heart rate are closely affected by shifting emotions. This combination of inputs through unobtrusive data collection can be paired with machine learning to classify and detect depressive tendencies in the user. The AutoMAP conceptual framework (Fig. 2.4) shows the proposed stages leading up to an improved ER approach focusing on older-aged people.

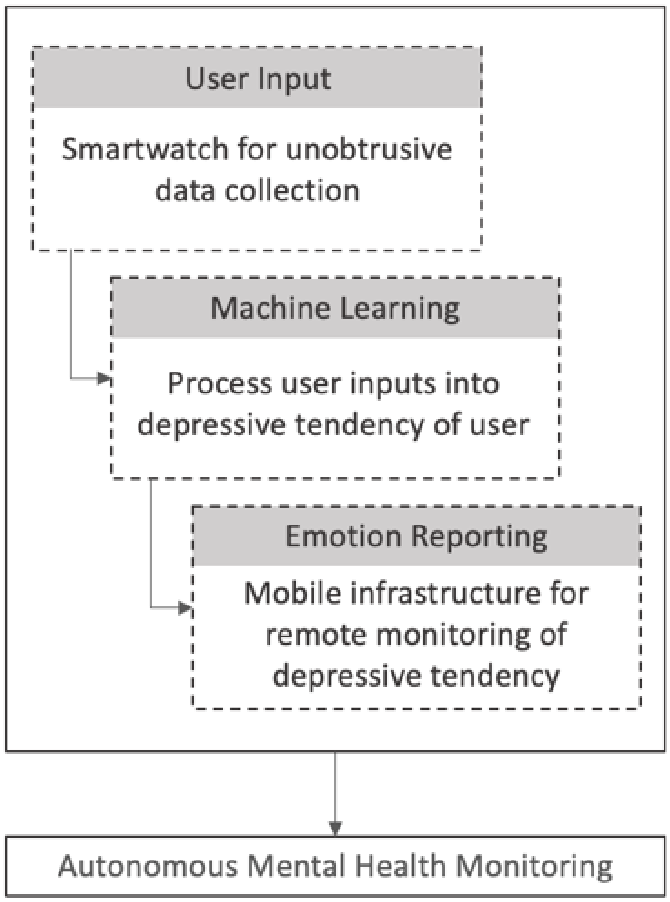



Figure 2.4: The AutoMAP framework

### 2.5.1 User Input

As a major focus of the current study is minimal user interaction, data needs to be collected easily. From the review on recent emotion recognition studies, wearable devices are the most unobtrusive of all other approaches such as on-body chest monitors or EEG headsets. In the case of facial recognition, users need to be within the cameras frame of focus and speech recognition requires explicit user interaction. Based on this, we have decided to use wearable smartwatches to collect physiological data that is most commonly altered with shifting emotions, i.e. walking, sleep and heart rate (J. A. Naslund et al., 2016; Zhang et al., 2016; Quiroz et al., 2018). It is expected that this combination of physiological factors appears will be able to classify emotions accurately and precisely as opposed to individual factors on their own.

### 2.5.2 Machine Learning

This component of the proposed framework serves as the 'brains' of the monitoring system which will process and classify the depressive tendencies of the user. The data collected from the smartwatch will be used to determine the ideal combination of factors for the highest accuracy of depressive tendency classification, e.g., gait + sleep, gait + heart rate, heart rate + sleep, gait + heart rate + sleep. Collected data will be pre-processed and then used to train a classification or regression model to classify or score a person's mental health level. Following this, the performance accuracy in the final validation stage will be performed.

### 2.5.3 Emotion Reporting

For reporting, a mobile app will be used to provide updates or notify caregivers of the user's mental health status, depression, or depressive tendencies. In the future, this may be provided to caregivers, staff, or family members. The app will mainly focus on reporting the user's depression levels, however since the user's physiological data will also be stored, it is likely the app will include changes in heart rate, sleep patterns as well as gait information. To avoid subconscious emotion alteration, it is advisable to avoid providing real-time emotion updates to the user themselves.

## 2.6 Conclusion

As analysed by the review, some of the more common approaches being taken currently include Electroencephalography, Facial Recognition, Speech Recognition, Voice Recognition, Heart Rate Variability, Electro Dermal Activity or Galvanic Skin Response, Respiration, Skin Temperature, and Electromyography. While the study highlighted various methods being used for emotion recognition, it was also stated that different approaches could be more suitable for certain studies based on application area. Some methods perform better with the use of stand-alone sensors, while others are more efficient and accurate with a combination of sensors and collected data. Data gathered by physiological sensors contains noise that can be moderated or completely eliminated in a lab-based setup resulting in a reduction in performance accuracy outside of a lab-based setup. Most real-world means of gathering emotion data using smart wearables, such as smartwatches, collect data over a long time period; while most studies on using smartwatches , focus on shorter time periods like minutes or seconds. This shows that there is still a gap in emotion data collected over a long time period and real-world emotion recognition. In comparison, routine-based approaches lead to easier, more intuitive, and transparent data collection and interpretation (Kasteren et al., 2017). Additionally, most studies to date have had very few test subjects, where classifier performance is relatively poor than if a larger sample size was used.

Many of the more common ER methods focus on outward expression, such as facial recognition, speech recognition or hybrid audio-video recognition (De Silva et al., 1997).

We must consider that a person who is happy might not be smiling, and a person who is sad might not be frowning, and could even be smiling. To address this gap, there needs to be further focus toward a potential solution that could analyse and perceive human emotions with implicit, i.e., minimal user input or interaction. The implications of age based cognitive impairment, as well as additional stressors like the inability to verbally express emotion or being bedridden, need to be accounted for.

Moving forward with the conceptual framework, we will be investigating its uses with a combination of physiological factors that are most affected by age related cognitive decline to detect and recognize user emotions, particularly negative emotions like depression. The aim is to use minimally intrusive means of data gathering from the user. With this approach, the goal is to propose a more autonomous emotion recognition-based monitoring system for older-aged people. Further, the aim of this research is to conduct feasibility and performance evaluations for the User Input and Machine Learning components of the framework.

The next chapter details the methods and materials that will be incorporated within this research.

# Chapter 3

## Materials and Methods

This chapter offers detail into the various components that were encompassed in the study design. Each section details the methods and materials involved in data procurement and analysis. These components come together to develop and validate the AutoMAP framework. This chapter also links how each design component addresses respective research objectives proposed in Chapter 1.

The AutoMAP framework in this research consists of three main components that correspond to the research objectives, namely (1) User Input (RQ1), (2) Machine Learning (RQ2), and (3) Emotion Reporting (RQ2 and RQ3). This research involves a single experiment, with multiple phases of analyses in a mixed design approach. This study was approved by the Human Research Ethics Committee (HREC) of University of Technology Sydney (HREC reference ETH20-4912).

## 3.1 Introduction

In 2021, 929 million people use smart wearables and 31 million use Fitbit devices, worldwide. While there is growing research on using smart wearables to benefit physical health, more research is required on the application and feasibility of using these devices for mental health and wellbeing. In studies focusing on emotion recognition, inference is often dependent on external cues, which may not always be representative of genuine inner emotion.

Cost effective smart wearables are increasingly used in the daily lives of the general population (WHO, 2017). As of 2021, there are 929 million smart wearable users around the world (Laricchia, 2019). Of these, approximately 31 million are Fitbit users (Curry, 2020). Consequently, there is a growing number of studies that are incorporating smart wearables, particularly Fitbit devices in health research (Bai et al., 2021; Chum et al., 2017; Ringeval et al., 2020). A recent scoping review investigated the effectiveness and efficiency of mobile health procedures for physical health (Domin et al., 2021). Over a 10-year period, including 148 studies, there was no 'one-size-fits-all' procedure for physical health. However, the authors found that mobile health interventions do exhibit promising effects for behavioral change. A similar review (Ridgers et al., 2016) found the need for more research on the effectiveness and feasibility of smart wearables for changing/assessing physical health. Incorporating Fitbit as the focal wearable device, Ringeval (2020) assessed the effectiveness of using Fitbit devices in interventions to promote and encourage healthier lifestyle outcomes. Evaluating 41 studies (Ringeval et al., 2020), it was concluded that Fitbit devices have the potential to improve lifestyle habits among users. Significant increases in daily step count, physical activity, and weight reduction were identified (Ringeval et al., 2020). A further review (Soon et al., 2020) investigated various studies to determine the applicability of wearable devices for outpatient vital sign monitoring. While concluding more clinical trials are required to investigate their validity and reliability, they found that early detection of physiological deterioration via wearable devices likely has a positive influence on patient outcomes. The studies assessed included on-body, potentially obtrusive sensors, such as heart rate monitors, patches, and arm bands. These sensors are effective in monitoring user vital signs, though they may not be feasible for prolonged use in everyday living.

The aforementioned reviews strengthen the idea that wearable devices elicit positive physical activity changes, however few studies have investigated how a person's mental health can benefit from the use of smart wearables (Chum, 2017; Luxton et al., n.d; Naslund et al., 2014). In addition, most of these studies used younger study participants, with older aged people (65+ years) excluded. User requirement studies and scoping reviews have been conducted around older aged people (Koo  Vizer, 2019; Elers et al., 2018, C. Chen et al., 2021), albeit it was highlighted that many factors are still not taken into account such as cost, usability barriers, security, privacy (Elers et al., 2018). In the case of patients with dementia, factors such as agency and self-esteem are not taken into account (Koo  Vizer, 2019). As of 2021, there are 264 million people of all ages worldwide suffering from depression (Hartmann et al., 2019). While the majority of older aged people are not depressed, studies have shown that older aged people are at a higher risk (WHO, 2017). Eighty percent of older aged people have at least one chronic illness, which can contribute to mental illness (CDC, 2021; NIA, 2017). Thirty percent of older aged people in residential care are at an increased risk of depression (Beyond Blue, n.d.). Mental illness in older aged people is commonly viewed as an inevitable reaction to or as a byproduct of changes in socioeconomic standing (WHO, 2017) or age discrimination (Han  Richardson, 2015) and as such, tends to be perceived as untreatable (CDC, 2021). Older aged people are also more likely to be concerned about the stigma of seeking treatment (Ridgers et al., 2016). Consequently, older aged people do not seek help when they feel depressed (CDC, 2021).

## 3.2 Experiment Design

This observational cohort study took place in zero-contact settings to accommodate COVID-19. We used smartwatches to collect user gait, sleep and heart data to derive their emotional state. Twelve participants consented and were sent participant packs which included a (1) Fitbit Alta HR device (Fig. 3.1), (2) participation confirmation, (3) distress management resource list, (4) simplified Fitbit user guide and (5) summarized task checklist. Participant pack documents can be viewed in Appendix A. Participants were asked to wear the Fitbit smartwatch over a 4-week period. Alongside this, they were required to fill out a daily questionnaire consisting of (1) a self-report mood rating Likert scale, (2) open questions on activity and food preferences to reduce response bias by diverting attention away from depressive symptoms, and (3) an optional section to add details of any out of the ordinary events in the preceding 24-hour period to account for any outliers in the collected raw data.


Figure 3.1: Chosen study device (Fitbit Alta HR)

Participants were also asked to fill out the Geriatric Depression Scale (GDS) (Greenberg, n.d.; Sheikh & Yesavage, 1986) once per week, as it was intended when initially designed and validated. The GDS was chosen as it is a global measure depression catering to older aged people and is also used by clinicians to provide insight on the participants' current emotional state. This is not a diagnostic tool; instead, it is intended to provide an overview of where the person falls within a range of depressive tendency scores. Figure 3.2 illustrates an overview of the study design.

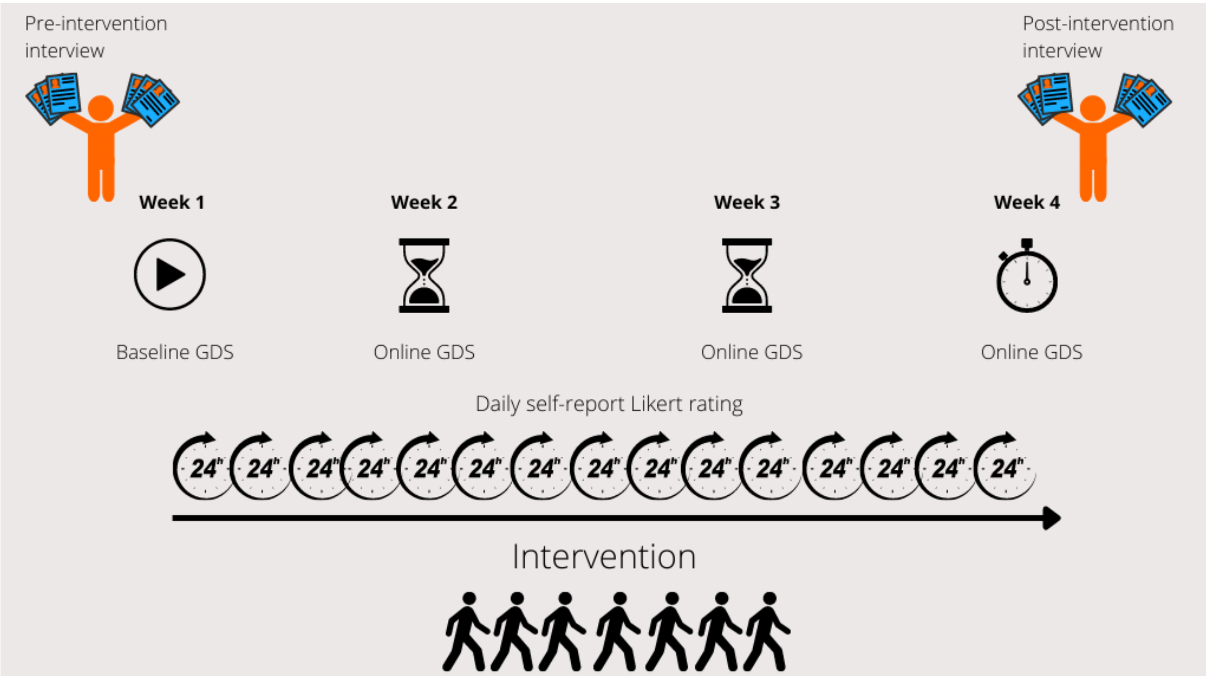



Figure 3.2: Intervention study design

## 3.3 Primary Objectives and Hypotheses

This study consists of multiple objectives and hypotheses, which are detailed in Chapter 1. In this section, the primary objectives and hypotheses are laid out to highlight the essential focus of this study.

**Objective 1** - Identify the usability of the smart wearables in older aged study participants.

**Hypothesis 1** - Users will not experience inconvenience in using the smart wearable in a routine-based approach (Kasteren et al., 2017), and as such will be able to incorporate it in their daily lives with ease.

**Objective 2** - Assess the correlations between physiological activity data and depressive tendencies.

**Hypothesis 2** - Participants with higher depressive tendencies (GDS scores) will exhibit differing physiological patterns than those will lower to no tendencies.

**Objective 3** - Determine whether physiological data can infer depressive tendencies in a user

**Hypothesis 3** - Heart rate, sleep and step counts combined will be indicative of depressive tendencies in a user, and effectively can be used to train a predictive model.

## 3.4 Participants

### 3.4.1 Screening

A call for expressions of interest (EOI) for participation in this study was circulated through emails and word of mouth through an internal database of older aged people interested in participation in research around healthy aging. Sixteen EOIs were received, and details were screened further for selection. Participants that were unable to provide written consent or were not contactable following their initial EOI were excluded from the study. Due to the small sample size, 2 participants aged 64 were also included, however they did not return for the post-study interviews. Their device data and pre-study interview data were excluded from all analysis stages of the study.

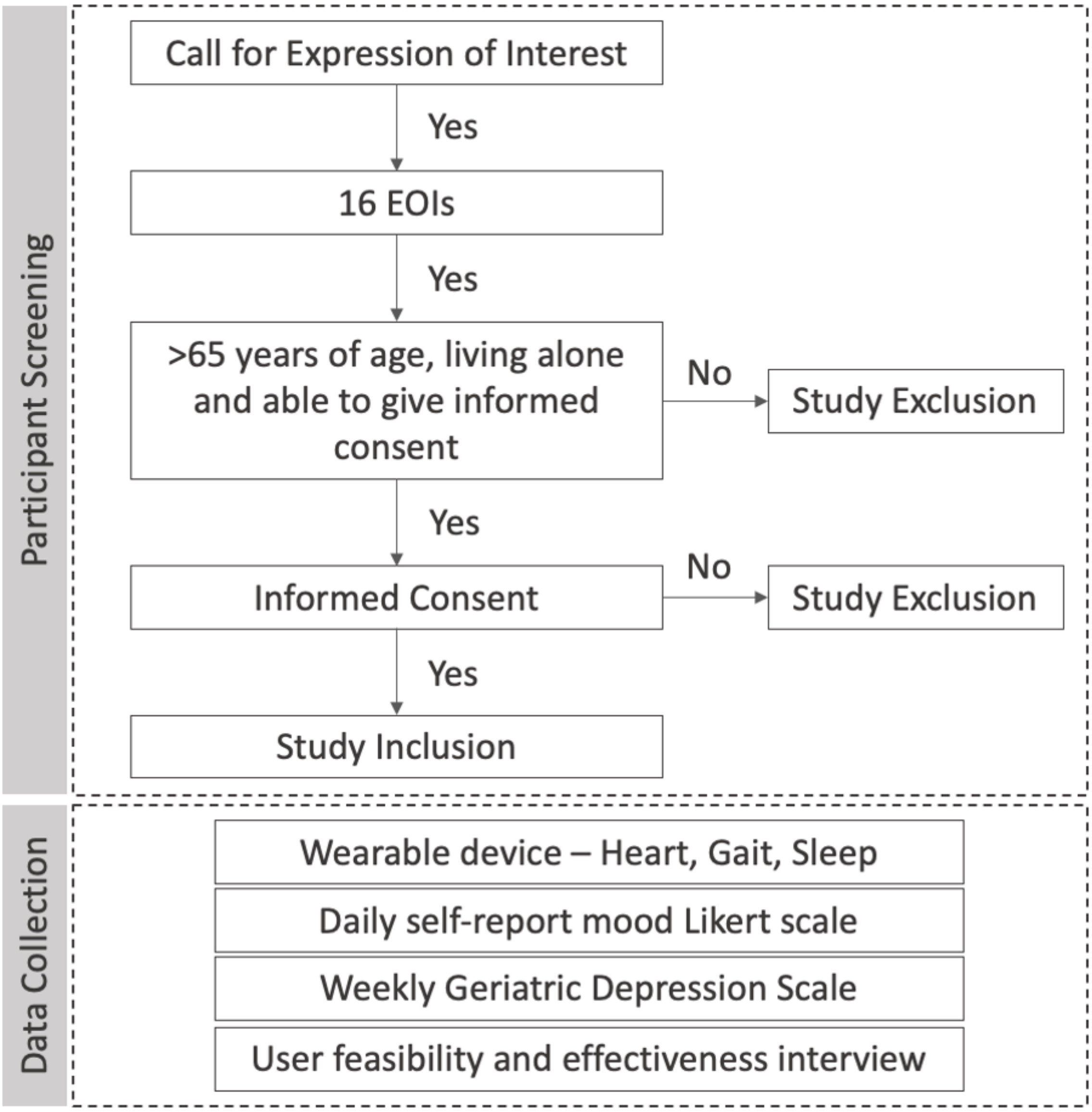



Figure 3.3: Participant recruitment and intervention workflow

### 3.4.2 Recruitment and Onboarding

English speaking people over 65 years old were eligible for preliminary inclusion in the study (Fig. 3.3). The inclusion criteria for the study included (1) being over the age of 65, (2) being willing to wear the Fitbit throughout the 4-week period, (3) living independently with no external assistance, (4) being able to give informed consent, and (5) having access to a computer, mobile phone and internet.

Independent living older aged people were chosen for this study as this is our target cohort for the envisioned autonomous mental health monitoring implementation of the AutoMAP framework. A summary of all inclusion and exclusion criteria is shown in Table 3.1. At this preliminary stage we explicitly excluded participants with pre-existing mental health issues, as we are in the exploratory stage of this research. Refer to Section 3.6.2 for further detail on the mental health state screening protocol.

After screening the initial expressions of interest (Fig. 3.3), participants were provided with an information sheet detailing the inclusion criterion and requirements during the 4-week period. Twelve participants were sent a participant pack that included (1) a Fitbit Alta HR device, (2) participation confirmation and information sheet, (3) distress management resource list, (4) simplified Fitbit user guide and (5) summarized task checklist. Using G*Power, sample size was calculated prior to the study with an expected p-value of 0.05. The critical F-value showed that a minimum sample size of 8 people would be sufficient. This sample size is required for future analysis on the collected smartwatch data and its use in training an initial test classifier for emotion recognition.

### 3.4.3 Participant Protocols

Participants were provided with a list of resources (Appendix A.8) available to them in case of any emotional distress during the study period. While the introductory interviews were basic questions about the everyday lives of the participants, we set protocols to terminate the interview session and stop participation if a participant got distressed at any point, especially during baseline GDS entries. Refer to the Interviews section for more details on the GDS.

Table 3.1: Inclusion and exclusion criteria

| Inclusion Criteria | Exclusion Criteria |
|---|---|
| Be above the age of 65. | Have pre-existing conditions affecting their sleep, walk or heart readings. |
| Be willing to wear the Fitbit throughout the 4-week period. | Have travel plans within the duration of the 4-week period. |
| Be living on their own. | Are unable to walk. |
| Be able to give informed consent to participate in the study. | Foresee out of the ordinary plans within the duration of the 4-week period. |
| Have a computer, mobile phone, and internet access. | Are unable to access online surveys as per the time requirement. |
| Live alone and do not require external assistance. | Have a history of skin rashes around their wrist area. |
| | Have a nickel allergy that may cause a reaction to the smartwatch charging probe. |

### 3.4.4 Researcher Interventions

Researchers in the study could advise the participants to seek assistance from a provided resource list or medical professionals. Participants were provided with point of contact information for the research team.

### 3.4.5 COVID-19 Protocols

This study took place in zero-contact settings to accommodate for COVID-19 protocols. Participant document packs and devices were sent via post and all interviews were held via phone.

## 3.5 Data Collection

### 3.5.1 Materials

Data collection was performed using a) a Fitbit Alta HR, b) the GDS questionnaire, and c) a 0 -10 self-report mood Likert scale. Participants were provided with onboarding packs upon confirmation of study participation. Along with the Fitbit Alta HR device, these onboarding packs included a (1) participant information sheet (PIS), (2) consent form (CF), and (3) reminder checklist.

This research project incorporates self-report mood Likert scale responses, geriatric depression scale (GDS) scores, step count, sleep patterns, and heart rate data from a smart wearable device for the quantitative portion of the study. The qualitative portion of this study consists of pre-post intervention impact and feasibility interviews conducted with participants directly before and after the study period.

Data extracted from the Fitbit device will be used to determine (a) whether participants found the device convenient and unobtrusive, and (b) the correlations between physiological data and depressive tendencies in a user. Data labelling and model training will be detailed further in Chapters 5 and 6.

To determine the feasibility and applicability of the Fitbit Alta HR for the User Input component of the AutoMAP framework, participants will be interviewed pre- and post-intervention. These semi-structured interviews will undergo thematic analysis, which will be detailed and discussed further in Chapter 4. Post-intervention interviews will also aim to gather feedback for the Emotion Reporting component of the framework.

### 3.5.2 Wearable Device

For this study, all participants were provided a smartwatch. Based on literature(Drew, 2018), we chose the Fitbit Alta HR smartwatch for our study due to its cost effectiveness for the eventual end user, as well as ease of setup and usage. Studies have shown fair performance accuracy for similar Fitbit devices (Fitbit Charge 2), with a heart rate estimation error of 14% (Bai et al., 2021). A review of Fitbit centred sleep studies additionally showed sensitivity values of 0.95-0.96 and specificity values of 0.58-0.69 for detecting sleep stages (Cook et al., 2017). Participants were required to wear the device at all times during the 4-week period with the exception of showering. As the study required sleep data, the participants were asked to wear the device when sleeping. To ensure minimal loss of data, we advised the participants to leave the Fitbit to charge before they showered (approx. 30 minutes). However, as the study aimed to mitigate any inconvenience and work in a routine-based approach, participants were not strictly required to abide by set charging times.

## 3.6 Outcome Measures

### 3.6.1 Physiological Measures

Using Fitbit and custom Application Programming Interfaces (APIs), optical heart rate sensor data (heart rate zone minutes), heart rate sensor and motion sensor data (sleep zone minutes), and pedometer data (step count) will be extracted for a series of quantitative analyses. Preliminary, exploratory analyses will be conducted for correlation observations. User data will be labelled with their calculated GDS scores to train and evaluate the predictive machine learning model. Findings will be used to determine subsequent directions for predictive modelling, which will effectively align with the second and third research questions.

### 3.6.2 Questionnaires

During the introductory briefing, participants completed the GDS questionnaire (Greenberg, n.d.; Sheikh & Yesavage, 1986) which acted as baseline measurements. The GDS consists of 15 yes/no questions with a single point score for each response indicative of depression. Any participants scoring >10 would be advised consult a GP prior to participating in the study. Higher scoring participants were not excluded immediately, as the GDS is not a diagnostic tool for measuring depression.

## 3.7 Data Extraction

This research follows a mixed-design approach, implementing both, qualitative and quantitative analyses. Data collection for this study required a single experiment and was conducted over a 4-week period. Participants were interviewed pre- and post- study through semi-structured interviews. The first set of interviews were conducted upon starting the 4-week study, with the second set of interviews being held upon study completion. These interviews were intended to gain feedback on the applicability, acceptability, and feasibility of the user input and emotion reporting components in the AutoMAP framework. Both sets of interviews were transcribed verbatim by the primary investigator.

Participant GDS responses and self-reported mood Likert ratings were extracted as individual logs for each participant. Data were de-identified upon initial extraction. For this, participants were provided with participant IDs that were also used to create participant Fitbit accounts for the duration of the study. Interview data were also categorized by participant ID. Raw Fitbit data was extracted using custom APIs as well as readymade Fitbit APIs. Raw Fitbit data logs were also extracted for each individual participant. Heart rate minutes (H1 - Out of Range/Resting, H2 - Cardio, H3 - Peak, and

H4 - Fat Burn), sleep zone minutes (S1 - Deep, S2 - Light, S3 - REM, and S4 - Awake), sleep score, restlessness and daily step counts were extracted as daily summaries, in alignment with GDS and self-reported mood ratings. Preliminary data cleaning was performed to remove duplicate GDS and mood ratings, i.e., if a participant responded twice in the same day, the initial reading would be used. Following any necessary transposition and data exclusions, data were merged. Individual Fitbit measures, GDS scores and self-report mood ratings were collated into a large single dataset, with deidentified participant data. This dataset will be used for model training and evaluation. Interview data will be used to validate the feasibility of the AutoMAP framework.

## 3.8 Discussion

This research follows a mixed-design approach. Data collected during the 4-week study period is intended to provide insight on the feasibility and predictive performance of the components in the AutoMAP framework. As these were semi-structured interviews, emerging themes were then used to determine further directions to take during the analysis. To investigate the applications of Fitbit data in predicting depressive tendencies in a user, the combined dataset was analysed through a preliminary exploratory analysis. The aim was to assess for strong correlations between physiological data and GDS scores. More detailed information and stages of both analyses will be detailed in Chapters 4, 5 and 6. Figure 3.4 maps out a summary of the analyses and their respective stages leading up to aimed primary outcomes.

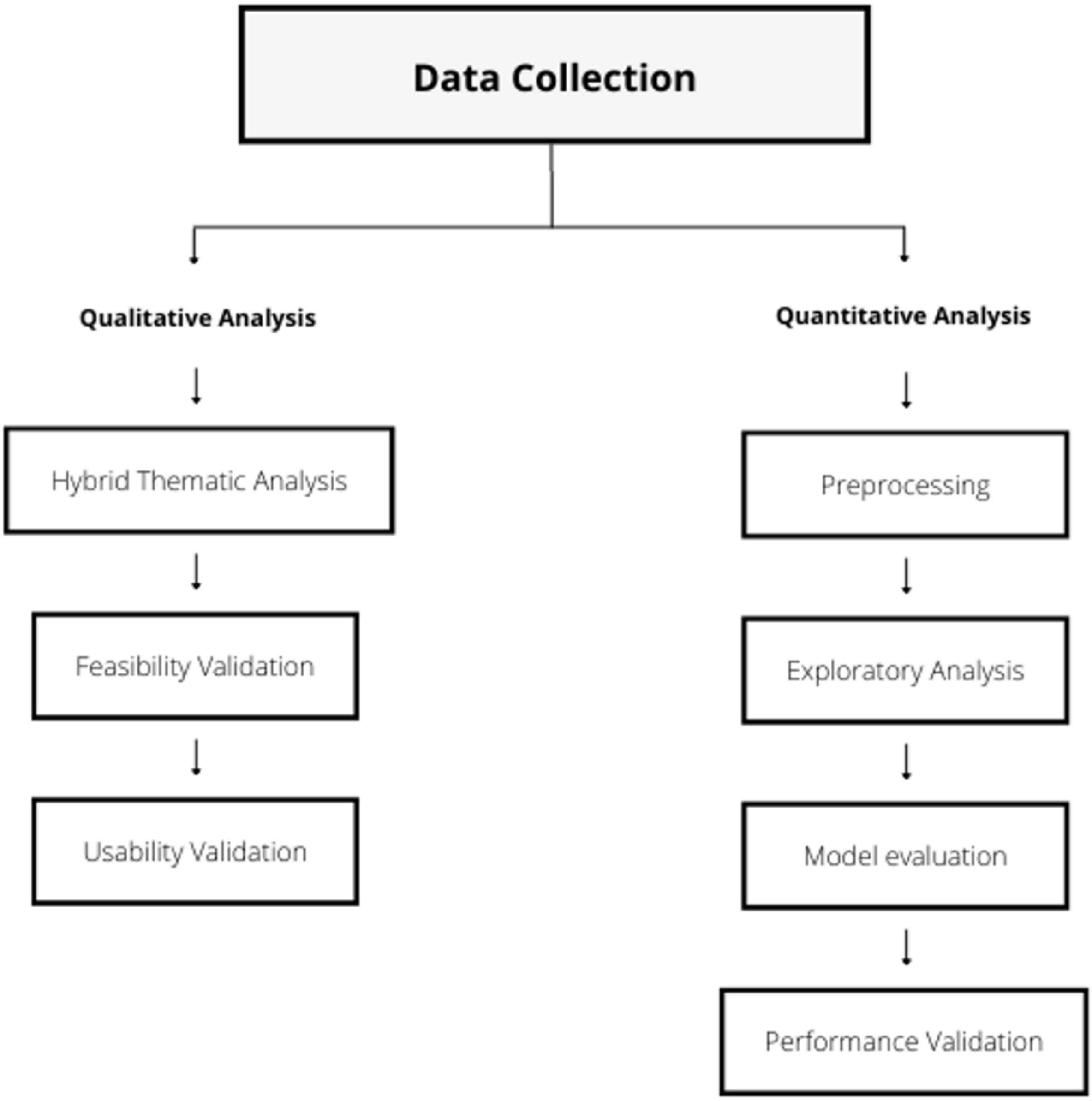



Figure 3.4: Mixed design data analysis stages and workflow.

## 3.9 Conclusion

This section presented the methods and materials used in this research. In line with the primary objectives of this research, these methods will allow for the validation of the AutoMAP framework. Individual components in the framework are assessed and evaluated for their feasibility, usability, and applicability within the overall framework. The remainder of this thesis provides more detailed insights into the experiment, analyses, and findings relating to the objectives of this research.

The next chapter dives into the qualitative aspects of this research, that is, the evaluation of using smart wearbles for data collection, as well as the potential impact of the AutoMAP framework as a whole. Chapter 5 subsequently presents a descriptive overview of the physiological device data, with Chapter 6 focusing on model development and evaluation.

# Chapter 4

# Feasibility of using wearable devices as means to monitor older people in an unobtrusive manner

This chapter presents the outcomes of a feasibility evaluation, to determine whether the AutoMAP framework is usable, convenient and and potentially implementable in older aged people. Findings from this analysis stage will validate the feasibility and applicability of the framework.

## 4.1 Introduction

Treatment of mental health conditions can benefit from the use of smart wearables, particularly Fitbit-like devices. A recent study investigated the feasibility of mobile health technologies to increase physical activity among users with severe mental illness (Soon et al., 2020). Feasibility was evaluated through frequency of use and acceptability through follow-up interviews with study participants. Participants reported high satisfaction levels, and increased motivation through goal setting and self-monitoring. Another qualitative study similarly integrated Fitbit devices with behavioral activation therapy (Segal et al., 2005). Findings showed positive self-awareness, peer-based and self-goal setting motivation, while negative feedback consisted of inconvenience, disinterest, and inaccuracy. Furthermore, a 12-week randomized trial incorporated a non-Fitbit device (tablet) and telephone counselling to increase physical activity (Vaportzis et al., 2017). The intervention was feasible and accepted by the participants, and resulted in weight loss and increased physical activity time in older aged people. Focusing on mental health, a study assessed sleep quality among users with major depressive disorders (MDD) using Fitbit Flex devices (Cook et al., 2017). The study aimed to compare research grade Actiwatch-2 versus Fitbit Flex wearables to evaluate sleep quality. They reported that the Fitbit Flex overestimated sleep time and had higher inaccuracies in comparison to the Actiwatch-2.

Factors such as heart rate and gait patterns can be indicative of depression (No Isolation, 2018; Pollreisz & TaheriNejad, 2017). Additionally, research has validated the use of smart wearables for emotion recognition (Drew, 2018; J. A. Naslund et al., 2016; Quiroz et al., 2018; Zhang et al., 2016). These studies utilized user data extracted through wearable smartwatches or bracelets, paired with external emotion eliciting stimuli to evaluate the reliability and accuracy of their approaches for emotion recognition. In lab-based settings with short timeframes of less than a few hours, three studies provided participants with stimuli to elicit happy/neutral/sad emotions (J. A. Naslund et al., 2016; Quiroz et al., 2018; Zhang et al., 2016). Stimuli included video and music clips of happy/neutral/sad settings and post stimuli activities included walking with a chest-mounted heart rate monitor (Quiroz et al., 2018). Time series analysis and statistical modelling were used for emotion prediction, with results showing a higher accuracy (74%) when detecting happy emotions as opposed to sad emotions. An earlier study required participants to walk for one minute after receiving a visual stimulus, to observe walking behaviors corresponding to the elicited emotion (Zhang et al., 2016). The algorithm had an overall recognition accuracy of 81.2%. Adopting a different approach for emotion elicitation, one study consisted of participants taking a short break to reduce interference with the previous stimulus set (J. A. Naslund et al., 2016). Using electrodermal activity, skin temperature, heart rate and a Self-Assessment Manikin form, researchers (Pollreisz & TaheriNejad, 2017) developed an algorithm based on EDA signals. The algorithm achieved an accuracy of 57% for emotion tagging. These studies show that gait patterns can reflect emotions and smart wearables can be used to recognize user emotions.

Mobile sensing is increasingly being used for the detection of emotion states in users (Gu et al., 2017; Lyons et al., 2017; J. A. Naslund et al., 2016; Quiroz et al., 2018; Soon et al., 2020). Such studies often use single variables or are performed under controlled environments. This reduces the efficiency and applicability in the real-world, as the surrounding stimuli and setting may alter the accuracy and performance of lab-based outputs. One such study using passive smartphone data and baseline depressive symptoms was performed to test the accuracy and effectiveness in predicting daily user moods. They noted that passive smartphone data alone is unsuitable for daily mood prediction, due to subjective phone usage patterns and daily mood reports (Pratap et al., 2019). However, Rodriguez et al. (Servia-Rodriguez et al., 2017) highlighted that passive sensing data has a 70% accuracy in predicting user moods for behaviour tracking and potentially interventions. Looking into the accuracies of using smart devices for user data collection, a comparison of consumer-grade smart wearables was performed (Bai et al., 2021). The Fitbit Charge 2, Fitbit Alta, and the Apple Watch 2 were assessed for heart rate, step counts and moderate-to-vigorous activity minutes (MVPA). This analysis found that the Fitbit Alta outperformed other devices for step counts, while coming second to the Fitbit Charge for MVPA. All monitors performed fairly well overall, particularly for step count accuracy.

Sultana et al. (2020) noted that current assessments of emotional states rely heavily on self-reports. To this, they conducted an exploratory study on the feasibility of leveraging machine learning algorithms with smartphone and smartwatch sensor data to detect emotional states and transitions in users. The study found a promising association between everyday movements and a person's emotional state. Further, a study was conducted to monitor changes in severity of depression in patients through wearable and mobile sensors (Pedrelli et al., 2020). The study found that physiological data from mobile and wearable sensors were feasible and may produce fair estimates of changes in depressive symptom severities in users. A similar study by Rykov et al. (2021) used biomarkers such as steps, energy expenditure, sleep, and steps to determine the risk of depression in users. The study contended that digital biomarkers based on behavioral and physiological data from smartwearables can be beneficial in depression screening. Rykov et al. (2021) further highlighted the need for further investigation into the predictive capabilities of this data. This strengthens the need for more real-world settings in user centred studies. Many emotion recognition studies (Benaissa et al., 2018; Quiroz et al., 2018; Sivapalan, 2014; Zhang et al., 2016) are performed under lab-based or controlled settings with explicit stimuli administered for emotion elicitation. Saeb et al. (2015) determined the effectiveness of short-term contextual sensor data and found a strong relationship between features extracted from a number of sensors and user reported emotion states over a two-week timeframe. They highlighted that long-term sensor measurements are more suitable for depression evaluation, as opposed to short-term data or momentary mental state ratings.

In Chapter 2, we proposed autonomous mental health monitoring for older aged people (AutoMAP); an emotion recognition framework using minimal to no explicit user input or interaction (Fig. 3). This prototype would operate via smartwatch device data and physiological user inputs. In this paper, we focus on validating the feasibility and acceptability of the AutoMAP framework. Therefore, this paper presents the findings of a feasibility study. The efficacy and accuracy of the emotion recognition techniques in this framework are beyond the scope of this paper. The next section describes the methods we applied for study setup, procedure, and data analysis. We then present our results, followed by a discussion on our findings.

## 4.2 Methods

As explained in Chapter 3, our study aims to acquire insight on the feasibility and applicability of AutoMAP for older adults. Participants wore a Fitbit smartwatch for a 4-week period, completed a validated depressive scoring survey weekly, and self-report their mood daily.

To identify the feasibility, applicability, and practicality of our framework, we performed a hybrid inductive/deductive thematic analysis on pre-post procedure semi-structured interviews with participants. The study protocol is publicly accessible under Open Science Frameworks (Mughal, Raffe, Stubbs, & Garcia, 2021).

## 4.3 Data Acquisition and Analysis

### 4.3.1 Questionnaires

Participants completed a daily questionnaire consisting of (1) a self-report mood rating Likert scale, (2) open questions on activity and food preferences (diversion questions away from depressive symptoms), and (3) an optional section to add details of any out of the ordinary events in the preceding 24-hour period. For the mood scale, participants were asked 'How would you rate your mood in the past 24 hours', rated on a scale from

1-10 once per day - with 1 being the lowest mood and 10 being the best mood. Meanwhile, diversion questions, such as 'What do you feel like eating?' and 'What do you feel like doing?' were added to the daily mood reported questionnaire to reduce response bias.

The daily questionnaire was intended to be performed consistently at the same time every day and at a time of the participants own choosing. Participants were also asked to report any out of the ordinary events during the day, such as 'Was there any out of the ordinary activity that improved or worsened your mood through the day?'. We will use the latter to explore the potential relationship between positive or negative spikes in heart rate or gait and their responses.

Participants also completed an online 15-item Geriatric Depression Scale (GDS-15) once per week (Sheikh & Yesavage, 1986). This short form of the GDS consists of 15 yes/no questions with a single point score for each response indicative of depression. The participants also completed the GDS-15 questionnaire online during the introductory briefing. These responses were used as the baseline GDS to observe outliers (if any). Any participants scoring >10 would be advised to have a GP consultation prior to participating in the study. Higher scoring participants were not excluded immediately, as the GDS is a screening tool rather than a diagnostic tool for depression (Greenberg, n.d.; Sheikh & Yesavage, 1986). The questionnaires can be found in Appendix B.

### 4.3.2 Interviews

**Pre-intervention (Introductory) Interview**

1. How have you generally felt over the past four weeks?

2. How often do your caregivers visit you/you visit them (family, friends, or medical professionals)?

3. Was this the same/less/more prior to the COVID-19 pandemic?

4. Do you feel that your lifestyle or your daily life has been impacted by COVID-19 in anyway?

5. Do you feel more dependent on others now as opposed to a few years ago?

6. GDS Questionnaire

7. Do you have any questions for us?

Follow-up questions will be asked based on their responses to the above questions.

**Post-intervention (Closing) Interview**

1. How was your experience over the past four weeks?

2. Did anything cause you discomfort over this time?

3. How was wearing the Fitbit for 4 weeks

4. Was the daily survey inconvenient?

5. Was there any part in the survey that made you feel uncomfortable or distressed?

6. Do you have any feedback or suggestions for how we can improve the study in the future?

7. Let's say we build a mobile app to send notifications to you or your families based on our findings, in real-time. Would that make you feel more relaxed while living independently?

8. What would you want in such an app?

9. Do you feel more at comfort when you know someone is looking after you or concerned about your inner wellbeing, regardless of whether they are with you all the time?

10. Do you have any further questions from us about our study?

11. Would your experience or answers with this study have been different prior to the COVID-19 pandemic?

Follow-up questions will be asked based on their responses to the above questions.

Figure 4.1: Pre-post interview questions.

Two sets of interviews (Fig. 4.1) were conducted with the participants; at baseline and at the end of the study period (week 4). Both sets of interviews were transcribed verbatim. Pre-procedure interviews were performed to provide insight into the everyday lives of participants in the 4-weeks preceding the procedure, as well as the impact of COVID on regular interactions (if any). Post-procedure interviews were performed to gain feedback on the procedure period itself.

The first set of semi-structured interviews (pre-procedure) were held in the introductory session, where participants also filled out the baseline GDS. The introductory interviews consisted of a series of questions relating to the everyday lives of participants and whether they were impacted by the COVID-19 pandemic. We added the COVID-19 component to account for anomalies in participant data, if they did feel impacted by the pandemic. The second set of semi-structured interviews (post-procedure) were held after the 4-week study period. Participants were provided more detail on the study and its objectives. To determine feasibility and usability, we asked the participants (1) about their experience during the 4-week period, (2) whether they would be interested in the autonomous mental health monitoring system we have proposed, (3) and what they would or wouldn't want to see in the mobile application. Such an application would report their emotional state to their caregivers. For example, clean, straight forward notifications versus a detailed dashboard view of vital signs. To gather insight on whether participants felt any inconvenience or discomfort during the 4-week study period, we asked participants a set of questions about the Fitbit device. Questions 1 to 3 (Fig. 4.1) assessed the user experience. Questions 4 - 6 and 10 were designed to obtain feedback for improvement on the study design, should the study be replicated in the future. Questions 7 and 8 asked participants about the likelihood of them being relaxed, given their caregivers are informed of their mental health remotely. Finally, question 9 was designed to determine the potential impact the framework as a whole may have on older aged people living alone.

We performed a hybrid inductive/deductive thematic analysis (Xu & Zammit, 2020) with four coders. Our deductive thematic analysis consisted of going into the analysis with an existing direction, concept, or theme. The analysis was initially directed toward acceptability and feasibility, however we developed sub-themes based on common texts to provide more insight and direction, as is typically performed in an inductive thematic analysis. In a workshop discussion setting among the authors, de-identified interview responses were assessed to (1) generate initial codes, (2) search for themes, (3) review the themes, and finally (4) define and name the themes. Both sets of interviews ran for approximately 11 minutes and were administered via phone call by the primary researcher and coder. The shortest interview ran for 4 minutes, with the longest running for 16 minutes. Figures 4.2 and 4.3 show an overview of how interview responses were coded and analyzed.

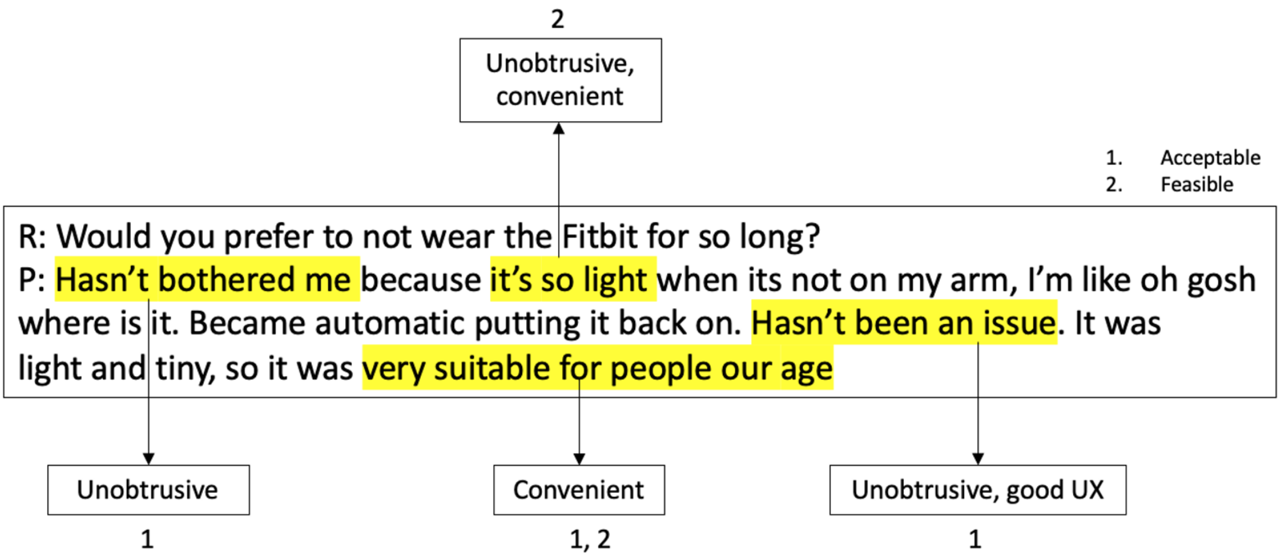



Figure 4.2: Sample excerpt analysis between researcher and participant (R-P) (UX = User Experience)

## 4.4 Results

A summary of participant demographics is presented in Table 5. Twelve participants (mean age = 68.58) were interviewed prior to the 4 weeks with 9/12 participants (mean age = 69.11) returning for a post-procedure interview (75% retention). The 3 participants that did not attend the closing interviews were uncontactable, and their Fitbit data and pre-procedure interview responses were excluded from the study. Closing interviews were conducted with the remaining participants, to identify potential improvements to the proposed framework (Fig. 4.1) and the effectiveness and feasibility of an autonomous mental health monitoring approach in reducing caregiver burden and dependence. Figures 4.2 and 4.3 provide an example of how interviews were analysed. Themes were formed based on key words and overall response context.

Table 4.1: Participant demographics

| ID | Gender | Age | GDS Mean (Baseline) |
|---|---|---|---|
| 1[a] | F | 64 | 3.406 (3) |
| 2 | F | 68 | 0.464 (0) |
| 3 | F | 72 | 0 (0) |
| 4[a] | F | 64 | 7.071 (6) |
| 5 | M | 67 | 0.316 (0) |
| 6 | M | 70 | 0 (0) |
| 7 | F | 68 | 0.974 (1) |
| 8 | F | 75 | 6.310 (9) |
| 9 | M | 71 | 0.4 (0) |
| 10 | F | 65 | 0 (0) |
| 11[a] | F | 73 | 2 (2) |
| 12 | M | 66 | 0.447 (1) |

[a] Excluded from study. Did not return for post-intervention interview.

**R**: How was your experience over the past four weeks?
**P**: Great it was good to sit down and put the thoughts down

**R**: Did anything cause you discomfort over this time?
**P**: No no, not at all. It was interesting to see how much walking I've been doing

**R**: Would you prefer to not wear the Fitbit for so long (4 weeks)?
**P**: Hasn't bothered me because it's so light when it's not on my arm I'm like 'oh gosh where is it?'. Became automatic putting it back on. Hasn't been an issue. It was light and tiny, so it was very suitable for people our age

**R**: Was there any part in either survey that made you feel uncomfortable or distressed?
**P**: No, I enjoyed it really

**R**: Let's say we build a mobile app to send notifications to you or your families based on our findings, in real-time. Would that make you feel more relaxed while living independently?
**P**: Yes

**R**: What would you want in such an app?
**P**: Vital signs, blood pressure (researcher note: semi-detailed)

**R**: Do you feel more at comfort when you know someone is looking after you or concerned about your inner wellbeing, regardless of whether they are with you all the time?
**P**: Definitely don't have an issue on that part

**R**: How was your experience over the past four weeks?
**P**: Was good. Didn't notice anything except one day got late with answers. Never an imposition and went very quickly. Didn't feel stressed, normal life.

**R**: Did anything cause you discomfort over this time?
**P**: None at all whatsoever. Actually surprised how quick it ended.

**R**: Would you prefer to not wear the Fitbit for so long (4 weeks)?
**P**: It doesn't worry me except call vibration (researcher note: minor technical difficulty)

**R**: Was there any part in either survey that made you feel uncomfortable or distressed?
**P**: None

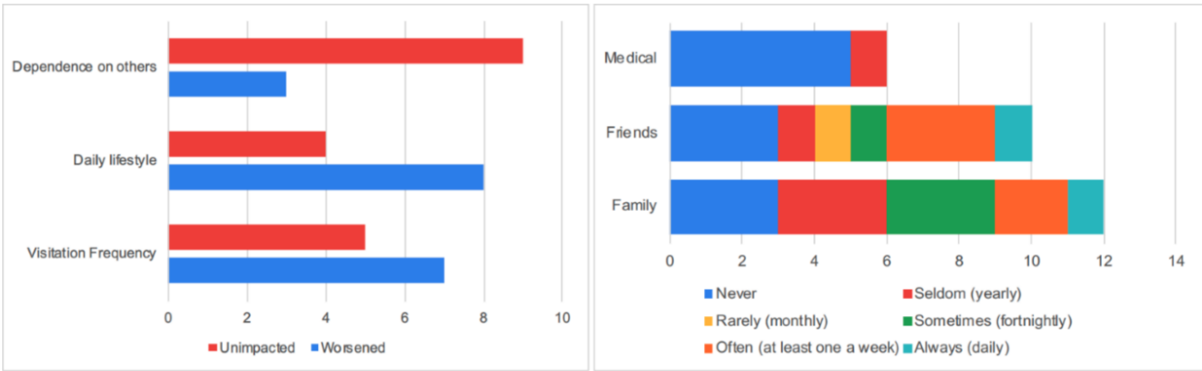
**R**: Let's say we build a mobile app to send notifications to you or your families based on our findings, in real-time. Would that make you feel more relaxed while living independently?
**P**: I think that would be nice but I don't have drops that major

**R**: What would you want in such an app?
**P**: A very clean app. Knowing physical details wouldn't particularly benefit anyone unless they have heart issues. I have a mentally disabled sister I think this would be good for. Should just say XYZ needs contact today.

**R**: Do you feel more at comfort when you know someone is looking after you or concerned about your inner wellbeing, regardless of whether they are with you all the time?
**P**: Yes that's a good idea

Figure 4.3: Excerpts from two researcher-participant (R-P) interviews with key words highlighted by one coder.

### 4.4.1 Pre-Procedure Analysis




*(a) COVID-19 Impact* *(b) Visitation frequency*

Figure 4.4: Pre-procedure themes: (a) COVID-19 impact on users in preceding 4-week timeframe. (b) Frequency of visitation with friends, family, or medical practitioners. Bars are different lengths as not all patients mentioned each of these items during the interviews.

Emergent themes from our analyses have shown that participants had a positive life outlook over the preceding 4-week period. Seventy-five percent of participants (9/12) had mostly positive emotions. Forty-one percent (5/12) experienced some negative emotions, of which 60% (3/5) were challenged with negative life events. Fifty percent of interviewees (6/12) expressed a range of health concerns, with 83% (5/6) showing minor concerns, and 17% (1/6) having major health concerns.

All 12 participants had family members visiting them occasionally, while 83.3% (10/12) had more frequent visitation with friends. Only 50% (6/12) had appointments with medical practitioners, usually 1×/ yr. Where participants were asked about visitation frequencies, responses (n=28) exceeded the interviewee count (n=12) as categories were non-mutually exclusive. i.e., participants could state visitations with any or all categories (friends/family/medical professionals). Categorically, 57% of the responses (16/28) showed visits ranging between family, friends, or medical practitioners 1 or 2×/yr. However, among participants that had more frequent interactions, 32% (9/28) indicated weekly interaction with friends. Of the 12 participants, seven people had less interaction since the start of the COVID-19 pandemic (Figure 4.4b). Additionally, 67% of the participants (8/12) felt that their lifestyle and daily lives were negatively impacted by the COVID-19 pandemic (Figure 4.4a).

## 4.5 Post-procedure Analysis

Responses after the 4-week study period were generally positive: 7 out of the 9 participants were pleased with the convenience and ease of study participation as well as the limited hands-on time commitment that was required. Figure 4.5 illustrates a summary of the thematic analysis for the post-procedure interviews. The user experience theme exceeds 9 participants, as 2 of the 9 participants mentioned this more than once through the interview. Moreover, 33% (3/9) of participants also experienced positive behavioral changes over the 4-week study period, particularly increased awareness and motivation to exercise more.

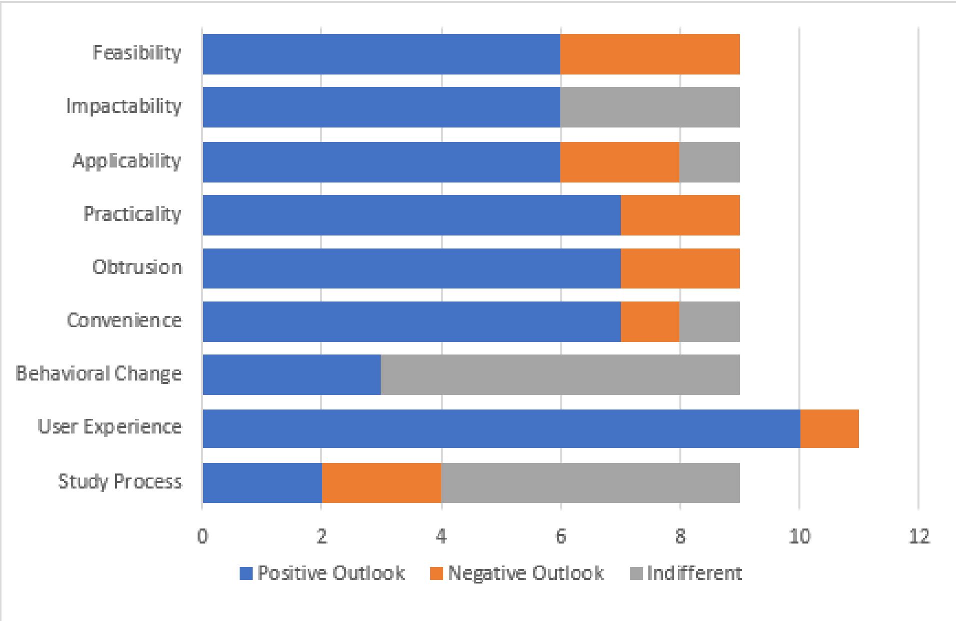



Figure 4.5: Dominant inductive post-procedure themes for overall autonomous mental health-monitoring system for older aged people (AutoMAP) implementation and overall procedure period. User experience exceeds interview count due to repeat emphasis in responses.

### 4.5.1 Study Feasibility

Six of the nine participants (67%) felt no discomfort during the protocol, while 1 participant felt mild discomfort with having to wear the watch during warmer summer days. Two of the nine participants (22%) stated feeling slight frustration and obtrusion with having to wear the device overnight. Seven of nine interviewees (78%) found no inconvenience during the procedure, with the most common concerns on compliance being with the required daily mood report timing. Two of the nine participants (22%) felt worried about responding correctly even though there were no right or wrong answers. One participant found the diversion question 'What do you feel like eating right now?' to be irrelevant, which was the purpose of the question. The procedure also resulted in behavioral change in 33% of the participants (3/9), particularly through exercise awareness and journaling during the daily self-reports. These participants mentioned feeling more motivated to increase their step counts, along with finding the open questions in the daily mood questionnaires to be therapeutic.

The device setup process and usage were favored by most participants, with only one participant having difficulty with the initial setup. Despite this, during data extraction, we found that two additional participants incorrectly linked their devices to the Fitbit application, which was resolved by requesting the participants to pair their devices again.

Overall, our analysis showed generally positive outcomes during the 4-week procedure, with a 75% retention. Some participants were seemingly aware of the study objectives, although not explicitly disclosed, while others found it to be a good aim once they were provided more detail on our study goals. Participants were given a high-level overview of the three components in the AutoMAP framework and how they are aimed toward autonomous mental health monitoring for older persons.

#### 4.5.1.1 Practicality & Applicability

Six of nine participants (67%) were interested in the full implementation of our working prototype and stated that they would also feel more relaxed and at comfort knowing that their mental wellbeing was being monitored, while the remainder felt it would be more useful for those people with major depression or diagnosed ailments. However, there were emergent themes pertaining to (1) privacy, (2) false positives (i.e., the end user is alerted when there is no problem), and (3) false negatives (i.e., the end user is not alerted when there is a problem). These will be addressed in the mobile application section. Twenty-two percent of the participants (2/9) suggested that the prototype would be more beneficial for users with specialized needs or diagnosed conditions such as Alzheimer's, Dementia, Autism Spectrum Disorder, Down Syndrome, or other known mental health issues.

One of the nine returning participants appeared to have a positive and promising on life during the pre-post procedure interviews, although their GDS scores alluded to a contradictory depressive state (mean = 6.310). One participant that did not return for the post-procedure interviews had higher GDS scores (mean = 7.071) and had an indifferent-to-positive outlook on life in the pre-procedure interview. Both participants also rated their own moods very highly (mean = 8.31 and 7.8) respectively, showing that self-perceptions of emotion are not always entirely accurate.

Seven (67%) of nine participants were pleased with the concept and potential of the prototype but raised concerns on its practicality. They suggested that the prototype may be a better fit for special needs persons of all ages or older aged people with more serious impairments, aside from general ageing populations. The envisioned end users of AutoMAP are the caregivers of older aged people using the smartwatch. However, one participant presented a circumstance where a user does not have any close friends or family and the end user could be missing or not concerned enough to follow up on alerts.

### 4.5.2 Mobile Application

The closing interviews also consisted of questions designed to collect suggestions and preferences for our proposed mobile application. The application would essentially notify users' caregivers of their depressive tendencies. While participants were not shown a visual mobile application, they were provided two design ideas for the application. One being a neat minimal application with only bare essential information such as their depressive score, and the other having more detailed information. Expanding on their preferences, fifty-six percent of the participants (5/9) preferred a clean, minimal application that would only send out plain-text alerts to the application end user (caregiver) when the device user's emotion levels reach a depressive tendency. For 2/9 participants, the application interface was preferred to be semi-detailed and more visually based as opposed to text-only. Another two participants did not specify interface preferences but raised concerns on potential issues such as misinterpretation or alarm in case of false positives, and reduced privacy. Addressing these concerns, participants would be allowed to choose what amount of information they want to be shared with their caregivers, with their depressive scores being the bare minimum.

## 4.6 Discussion

### 4.6.1 Principal Results

Despite some participants facing minor technical difficulties during the initial device setup and syncing process, participants were positive about the framework as a whole. The procedure resulted in positive behavioral change, not only physical wellbeing, but also mental wellbeing. Participants commented that the procedure made them more aware of their physical activity, while some took the daily survey as a means of journaling, which led to mental relaxation. Addressing concerns pertaining to the practicality and applicability of our prototype for those that have no caregivers of their own, we recommend a volunteer function within the AutoMAP mobile app that could allow other nominated persons to check on users. Living through the COVID-19 pandemic, this could also benefit the volunteers through added purpose or interaction.

Participants were concerned about the possibility of false positives sent to caregivers or the potential implication of false negatives. To mitigate this, we will need to train and test our algorithm to achieve high performance levels of emotion recognition from the smart watch data. Where privacy is a concern, device users might choose what extra information is visible to the caregiver. This could include (1) vital signs, (2) emotional range history, and (3) movement patterns. We need to consider whether device users should see their own emotion level information. This could cause unconscious bias or emotion alteration.

Although most participants favored a minimalist visual based notification application design approach, we will provide options for semi-detailed views instead of minimal or pure text-only for those that may prefer more information in the app. We propose a mobile application that (1) notifies caregivers when a user's scores are indicative of depressive tendencies, (2) facilitates autonomy and privacy, and (3) allows for interface selection. User scores will be determined through their physiological data and machine learning techniques.

### 4.6.2 Strengths and Limitations

This study has several strengths. The procedure was convenient and easy to implement from a user experience perspective. While some participants felt mildly concerned about the device setup and compliance with a consistent time for the weekly and daily surveys, all those who attended the follow up interviews were positive about the general study design and ease of participation. We blinded participants to our study aims during our pre-procedure interviews and informed participants of the study aims after the 4-week procedure. This type of blinding prevented subconscious response bias. The overall impression of the study was positive, with participants reporting that the 4-week study was well conducted, easy to follow and had no significant inconvenience to their everyday lives. Generally, the daily survey was perceived to be clear and concise, although in future iterations, efforts should be made to add more range to the mood rating components of the survey. This could potentially allow for finer, more detailed self-reported emotion mapping.

Our study analyzed participants over the age of 65, with no diagnosed mental health issues. A more diverse sample could improve the applicability and practicality of AutoMAP, as well as its performance and accuracy. While participants found the study favorable, some preferred to not wear the device overnight. There may also be issues pertaining to sharing private medical user data. The latter will be assessed and considered in future research. Despite our efforts to maintain engagement with all participants, we were unable to interview 3 of the 12 participants for post-procedure feedback. Including these dropouts in our analysis may have provided a different perspective on the feasibility and acceptability of the procedure and framework.

### 4.6.3 Comparison with Prior Work

Our findings align with previous research on smart wearables for emotion recognition (Quiroz et al., 2018; Shu et al., 2020) or mental wellbeing (Chum et al., 2017; Lyons et al., 2017). These findings show that wearable sensors can be utilized for emotion recognition. Some of these studies (Quiroz et al., 2018; Zhang et al., 2016) were in controlled settings, with deliberate emotion stimulation, where participants were not emotionally invested in the stimulus. These studies also observed participants for short time periods, meaning it is questionable whether the stimuli provided for emotion elicitation had the intended effect. By contrast, our procedure was performed in real-world settings, with participants keeping the Fitbit on during daily living. Therefore, collected sensor data reflects emotions in real day-to-day life.

Promoting behavioral change through telephone counseling and smart wearables has shown promising outcomes. Investigators also highlighted that competing wearable products may have different applications and accuracies, which should be accounted for in future replications of procedures using smart wearables. Another concern is the sustainability of the procedure. The aforementioned study used a device from the Jawbone company, which is no longer manufacturing devices. This could render the implementation of procedures using Jawbone devices potentially invalid for long-term implementation. In accordance with previous literature (Bai et al., 2021), we found Fitbit devices to be an acceptable tool to monitor user mental health, albeit a study on sleep among those with MDD found Fitbit Flex devices to be incomparable with research grade Activewathch-2. These findings however pertained to solely sleep focused analyses, while our focus lies on a combination of physiological inputs for depression detection. Additionally, our work differs in that we sought to determine the feasibility of Fitbit devices for preemptive depressive tendency detection, in contrast to implementing the devices for behavioral therapy for people with diagnosed depression.

## 4.7 Conclusion

On the basis of our pre-post interview findings, we will make modifications for future replication of our study, including (1) predefined weekly and daily survey times for all participants, (2) fewer diversion questions, (3) assisted device setup and walk-throughs and (4) more check-ins with participants during the study period.

Additionally, we gained valuable insight on end user preferences for the mobile application component of our framework. The combination of pre-post interviews enabled showed that everyday conversations and external cues alone are insufficient to detect depression or depressive tendencies in people alone. Future procedures should also recruit specialized cohorts with more diverse diagnoses, such as people with other diagnosed mental health issues such as dementia, autism spectrum disorder, and communication impairment. Special cohorts were explicitly excluded from this preliminary study.

This study is a vital step to validate our framework and identify requirements for the development of our proposed mobile application. Moving forward, we will develop the mobile application, as well as train and test our machine learning algorithm to detect depressive tendencies using physiological inputs. The machine learning component will be integrated with the mobile application, to notify caregivers if the user's score is indicative of potential depression.

The next chapter presents a descriptive overview of the extracted Fitbit data. The data is also used to develop a dataset with participant GDS score and self-report mood ratings. The resultant dataset is consequently used to train and evaluate the machine learning component of the AutoMAP framework.

# Chapter 5

# Exploratory Overview and Behavioural Insights

This chapter provides a descriptive overview of the extracted device data. Observations from this exploratory analysis will provide supporting evidence toward using physiological data to detect depressive tendencies in users. The resultant dataset will also be used to train predictive models for depressive tendency detection. A detailed look into the dataset also aims to visually investigate trends between participants with low, mild, or high depressive scores.

## 5.1 Background

Studies using smart wearables for emotion recognition achieved high accuracies >70% for binary classifications across happy vs neutral vs sad or happy vs neutral vs angry. Using smart watches and chest-mounted heart rate monitors for happy-neutral-sad classification, researchers achieved median accuracies of 78% (Quiroz et al., 2018). A similar study evaluated classification accuracies between pairings of neutral-angry-happy emotions. The highest accuracy achieved was of 91.3% for a neutral vs angry binary classification (Zhang et al., 2016). However, these studies were performed in a younger cohort with emotion stimulation, and in lab-based settings (Quiroz et al., 2018; Zhang et al., 2016). Physiological signals are frequently used for emotion recognition such as smart watch heart rate data and prior research has observed the applicability of using smart wearables along with accompanying sensors (Pollreisz & TaheriNejad, 2017; Quiroz et al., 2018; Shu et al., 2018; Zhang et al., 2016). An algorithm based on electrodermal activity, skin temperature, heart rate and a Self-Assessment Manikin form, achieved an accuracy of 57% for emotion tagging (Pollreisz & TaheriNejad, 2017).

While studies validated the applications of smart wearables for emotion recognition, they were performed in semi lab-based settings, or required stimuli to illicit negative, neutral, or positive emotions. Stimuli consisted of audio-visual inputs, with participants walking for for a short distance post-stimulation (Quiroz et al., 2018; Shu et al., 2020; Zhang et al., 2016). Participants were provided with audio-visual stimuli through a variety of movie scenes and mood setting music clips such as euthanizing a pet or a funeral movie scene to illicit sad emotions. Happy emotions were triggered through comedic and positive stimuli such as Monty Python or Au Claire de Lune. Following emotion stimulation, participants were asked to walk for one minute (Zhang et al., 2016) or a distance of 250 meters (Quiroz et al., 2018), with start and stop times being recorded. Additionally, on-body chest-mounted sensors, that could be deemed obtrusive when worn for a prolonged time, have also been investigated (Quiroz et al., 2018). Statistical modelling and time series analysis for emotion prediction showed a 74% higher accuracy for happy versus sad emotions. In a similar study using heart rate and accelerometer (gait) data, participants walked for one minute post visual stimulation. This resulted in an overall prediction accuracy of 81.2% for binary emotion classification Zhang et al., 2016). These studies provide evidence that human physiology can be used for emotion recognition. While there have been various studies on emotion recognition (Pollreisz &

TaheriNejad, 2017; Quiroz et al., 2018; Shu et al., 2018; Zhang et al., 2016), older aged people (65+ years old), emotion recognition and the relationship between physiological movements has been less explored.

Fitbit data were collected over a 4-week period. Participants also completed an online GDS form weekly, as well as a self-reported mood Likert scale daily. Both online forms were advised to be filled out at the same time throughout the study period. Device data were extracted using Fitbit out-of-box and custom APIs. GDS and self-reported mood Likert ratings were extracted through an out-of-box function provided by the form hosting provider. All data were exported as CSV (Comma Separated Value) files.

## 5.2 Data Extraction

The GDS and self-report mood Likert questionnaires were completed by participants online throughout the study. Data from these questionnaires were downloaded for each participant individually. Data from the Fitbit were extracted using a custom API, along with raw Fitbit data logs. The next sections detail each of the Fitbit measures that were extracted, and at what granularity.

### 5.2.1 Heart Rate

Four heart rate zones were extracted at minute level granularity. These zones are determined by the Fitbit device based on the users resting heart rate. The extracted zones were Out of Range/Resting (max heart rate), Cardio (70 to 84 percent of your max heart rate), Peak (85 to 100 percent of your max heart rate), and Fat Burn (50 to 69 percent of your max heart rate). For our analysis, we labelled each of these zones as H1, H2, H3, and H4 respectively.

### 5.2.2 Sleep

Four sleep zones per night were also extracted at minute level granularity. These zones are determine based on breathing patterns and heart rate during the sleep period. Deep, Light, REM, and Awake were labelled as S1, S2, S3, and S4 respectively, for our analysis. In addition to this, we also extracted sleep quality and restlessness scores. These measures were extracted at per night granularity, as single score summaries.

### 5.2.3 Step Count

Pedometer data were extracted as 'per day' summaries.

### 5.2.4 GDS and self-reported mood rating

Participants completed the GDS once per week and the calculated score was applied to the entire preceding week. The GDS was designed to calculate depressive scores based on how the participant felt over the preceding week through 15 Yes/No questions.

Participants also completed a self-reported mood scale. We compared these ratings to verify whether self-perception and evaluation of own mood are reliable or not (Mughal et al., 2020). That is, some participants rated their mood consistently as >8 on a 1 to 10 scale with 10 being the most positive mood, while their GDS scores measured concurrently indicated high depressive tendencies. The higher the GDS score, the higher the depressive tendency. This is further detailed in Section 5.4.2.

Heart rate and sleep zones were categorized by combined minutes in each respective zone for that day. These measures were extracted and analysed to observe their correlation with depressive tendencies in the user. Combined daily summaries were labelled according to the corresponding daily mood ratings and weekly GDS scores. Overall, in the 4-week study period, we obtained 2,365 Fitbit data samples across 11 measures (H1, H2, H3, H4, S1, S2, S3, S4, sleep score, restlessness, and step count), 32 GDS scores, and 215 self-reported mood ratings.

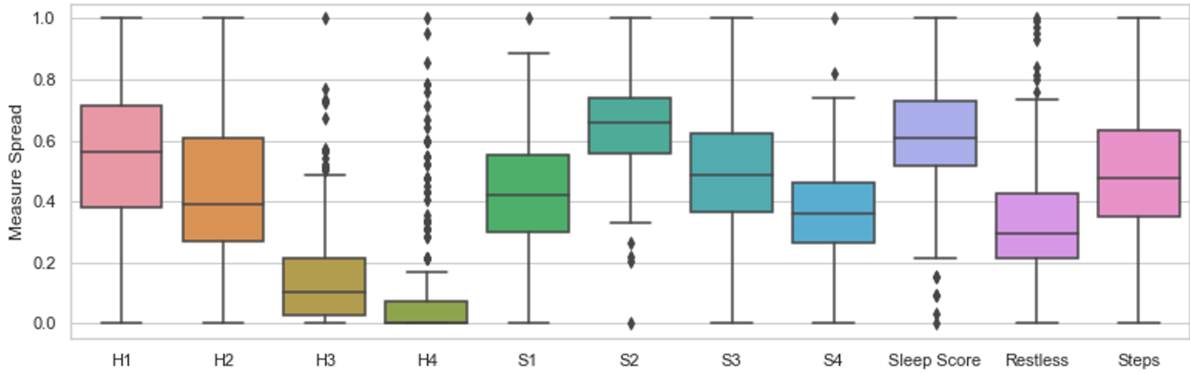



Figure 5.1: Activity similarities based on minute distribution spread.

## 5.3 Pre-processing

Extracted pre-post interview data, questionnaire responses, and Fitbit data samples were run through a pre-processing pipeline to transform the data for analysis. The data pre-processing pipeline consists of standard extract-transform-load (ETL) techniques (Kheirkhahan et al., 2019). Participant GDS responses and self-reported Likert ratings were extracted using features from the online form platform (Google Forms).

Raw Fitbit data were transformed into a readable data for analysis. First, raw Fitbit data were extracted using (1) a custom API, (2) Fitbit APIs, and (3) out of-box data logs. Heart rate data, sleep data and walking data were then transformed to output uniform data. For this, Fitbit data samples were normalized through rescaling within the range 0 to 1 (Fig. 5.4). Second, missing values (3 days ×1 participant) were imputed through replacement using means of non-missing values in the dataset. The data were then collated with participant mood ratings and labelled with participant GDS scores. These scores were calculated by the primary investigator using the validated GDS calculation guideline (Sheikh & Yesavage, 1986). The pre-processing ETL pipeline is visualized in Figure 5.2.

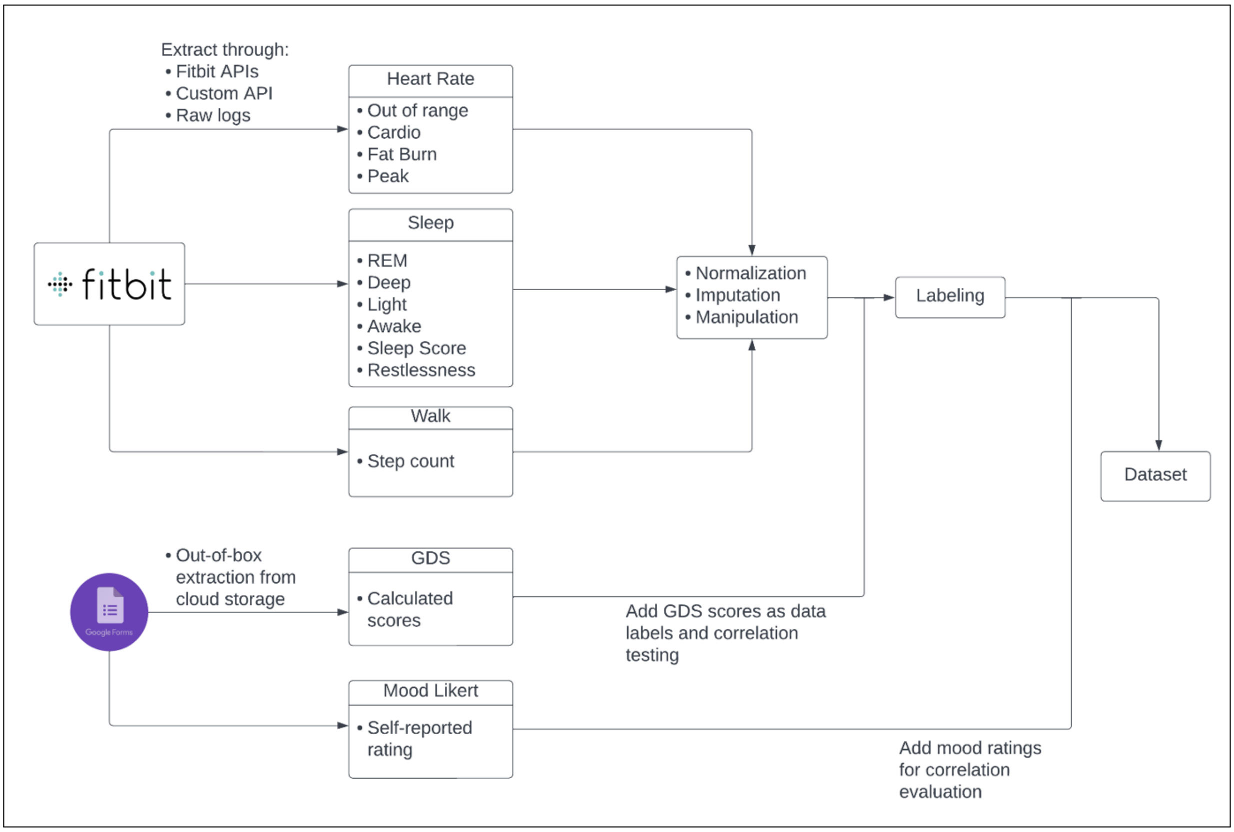



Figure 5.2: ETL pipeline: data extraction, transformation, and labelling for the resultant dataset.

### 5.3.1 Dataset Summary

The final dataset consisted of 215 samples. Only 8 / 12 participants were included in the dataset for further analysis. Following dataset preparation, preliminary exploratory analysis was performed. As participants were not recruited based on their baseline GDS scores, combined participant GDS data were right skewed. While data were normalized to scale all data points within a uniform range (Fig. 5.4), it did not have a significant impact on the skewness of the data.

## 5.4 Descriptive Analytics

| | H1 | H2 | H3 | H4 | S1 | S2 | S3 | S4 | Sleep Score | Restless | Steps |
|---|---|---|---|---|---|---|---|---|---|---|---|
| count | 215.000000 | 215.000000 | 215.000000 | 215.000000 | 215.000000 | 215.000000 | 215.000000 | 215.000000 | 215.000000 | 215.000000 | 215.000000 |
| mean | 749.102326 | 449.706977 | 57.413953 | 4.460465 | 56.586047 | 266.702326 | 81.381395 | 58.641860 | 77.767442 | 0.089377 | 9909.567442 |
| std | 276.958934 | 229.804290 | 64.987352 | 8.759351 | 23.058657 | 56.357792 | 25.858776 | 13.472267 | 6.473661 | 0.034583 | 4157.826963 |
| min | 2.000000 | 6.000000 | 0.000000 | 0.000000 | 5.000000 | 27.000000 | 12.000000 | 24.000000 | 58.000000 | 0.027837 | 8.000000 |
| 25% | 518.500000 | 283.500000 | 10.000000 | 0.000000 | 41.000000 | 233.500000 | 63.000000 | 49.000000 | 75.000000 | 0.066339 | 7074.500000 |
| 50% | 769.000000 | 412.000000 | 37.000000 | 0.000000 | 56.000000 | 271.000000 | 80.000000 | 58.000000 | 78.000000 | 0.081081 | 9693.000000 |
| 75% | 973.500000 | 640.500000 | 79.000000 | 3.000000 | 71.500000 | 301.000000 | 99.500000 | 68.000000 | 82.000000 | 0.105091 | 12871.000000 |
| max | 1364.000000 | 1048.000000 | 371.000000 | 42.000000 | 126.000000 | 398.000000 | 152.000000 | 119.000000 | 91.000000 | 0.209951 | 20293.000000 |

Figure 5.3: Raw data sample summary.

| | H1 | H2 | H3 | H4 | S1 | S2 | S3 | S4 | Sleep Score | Restless | Steps |
|---|---|---|---|---|---|---|---|---|---|---|---|
| count | 215.000000 | 215.000000 | 215.000000 | 215.000000 | 215.000000 | 215.000000 | 215.000000 | 215.000000 | 215.000000 | 215.000000 | 215.000000 |
| mean | 0.548533 | 0.425822 | 0.154755 | 0.106202 | 0.426331 | 0.646098 | 0.495581 | 0.364651 | 0.599013 | 0.337917 | 0.488123 |
| std | 0.203347 | 0.220542 | 0.175168 | 0.208556 | 0.190567 | 0.151908 | 0.184706 | 0.141813 | 0.196172 | 0.189896 | 0.204971 |
| min | 0.000000 | 0.000000 | 0.000000 | 0.000000 | 0.000000 | 0.000000 | 0.000000 | 0.000000 | 0.000000 | 0.000000 | 0.000000 |
| 25% | 0.379222 | 0.266315 | 0.026954 | 0.000000 | 0.297521 | 0.556604 | 0.364286 | 0.263158 | 0.515152 | 0.211418 | 0.348361 |
| 50% | 0.563142 | 0.389635 | 0.099730 | 0.000000 | 0.421488 | 0.657682 | 0.485714 | 0.357895 | 0.606061 | 0.292365 | 0.477446 |
| 75% | 0.713289 | 0.608925 | 0.212938 | 0.071429 | 0.549587 | 0.738544 | 0.625000 | 0.463158 | 0.727273 | 0.424205 | 0.634114 |
| max | 1.000000 | 1.000000 | 1.000000 | 1.000000 | 1.000000 | 1.000000 | 1.000000 | 1.000000 | 1.000000 | 1.000000 | 1.000000 |

Figure 5.4: Normalized sample summary.

### 5.4.1 Correlation Evaluation

To investigate the correlations between physiological outcome measures and GDS scores, correlation heatmaps were plotted using correlation coefficients. This coefficient represents the strength of the linear relationship between two variables (Schober et al., 2018). Due to the limited nature of the dataset, at this stage only linear relationships were investigated. Studies have been conducted to determine the reliability of small or limited datasets (Han et al., 2021; Torgyn et al., 2015). One particular study (Torgyn et al., 2015) contended that probability weighting, sampling design, and variable screening play a significant role in prediction performance, while stacking random forest can provide improvements over random forests. A negative coefficient indicates that the respective physiological feature changes in the opposite direction as a person's depressive standing. That is, if a person's GDS score or depressive tendency were to increase, a negatively correlated physiological feature would decrease. On the other hand, a positive correlation indicates changes in the same direction. Such as a decrease in depressive tendency resulting in a decrease in Cardio (H2).

Heatmaps were plotted with different variations of the dataset. These heatmaps are discussed further in Section 5.4.1. Individual participant heatmaps were visualized to observe patterns in correlation for low versus mild versus high depressive tendencies (based on GDS scores). Figures 5.5, 5.6, and 5.7 illustrate the correlations for three participants among low-mild-high depressive tendencies. The majority of the physiological features were negatively correlated with the GDS scores for the participant with low depressive tendencies, i.e., an increase in depressive tendencies for this participant resulted in reduced overall activity and reduced sleep quality.

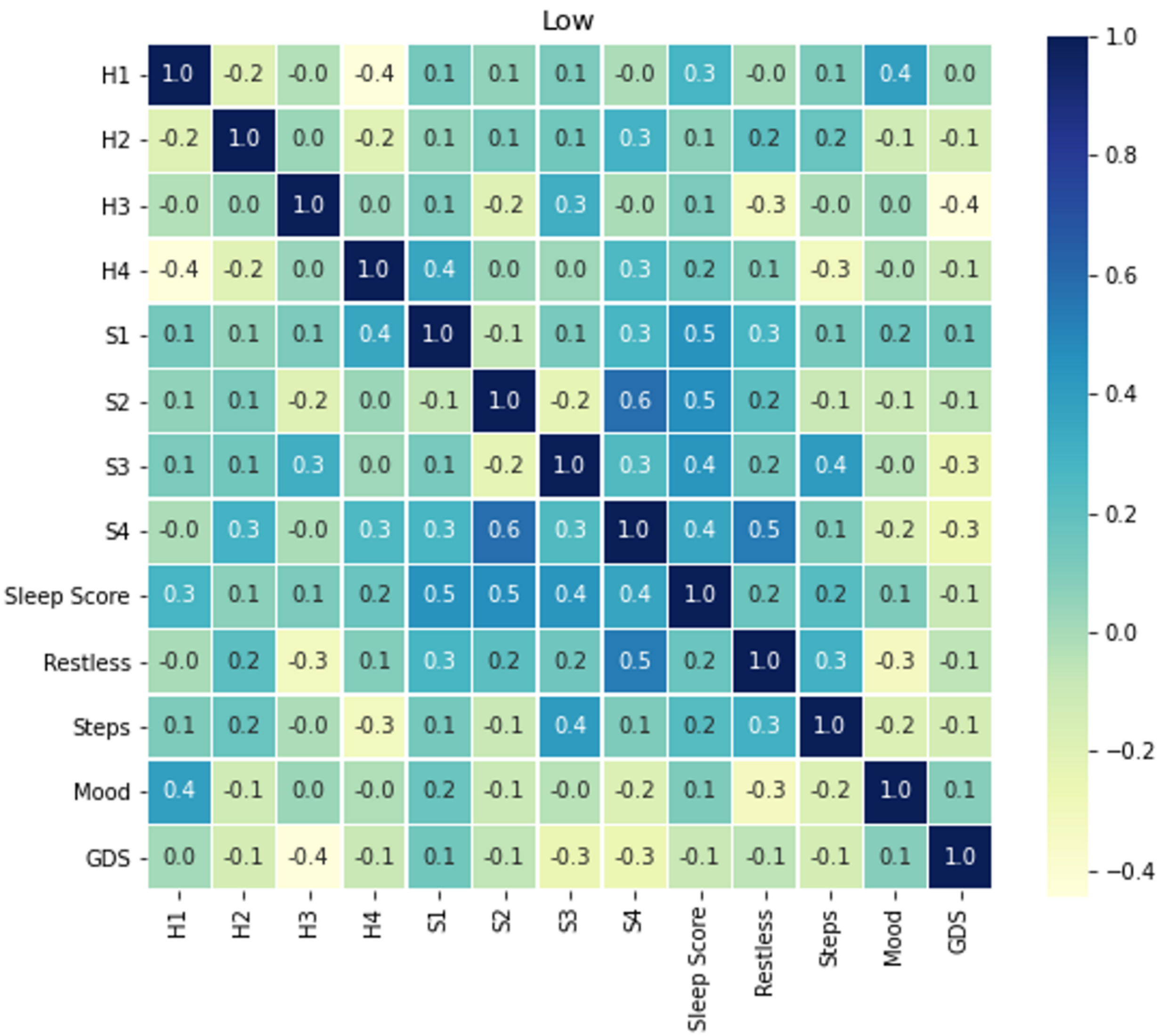



Figure 5.5: Correlations for low depressive tendency.

Fig. 5.6 shows a visualisation of the correlations for a participant with mild depressive tendencies. Heart rate feature correlations unlike with the participant with low depressive tendencies are more positively linked; i.e., the user has increased minutes of sleep in each zone when they have higher GDS scores. Similarly, sleep patterns for this participant are all positively linked to depressive tendencies (GDS) in the user, while the participant with low depressive tendencies had negatively linked sleep patterns. That is, with an increase in depressive tendencies, the user had reduced minutes in each sleep zone. Some features were correlated similarly between participants strengthening the idea that there is no 'one-size-fits-all' model that can be applied for mental health monitoring. Some people with depression may become more active, while others may become less active. This supports the use of multiple sensors to detect depressive tendencies in a user, as opposed to individual features on their own.

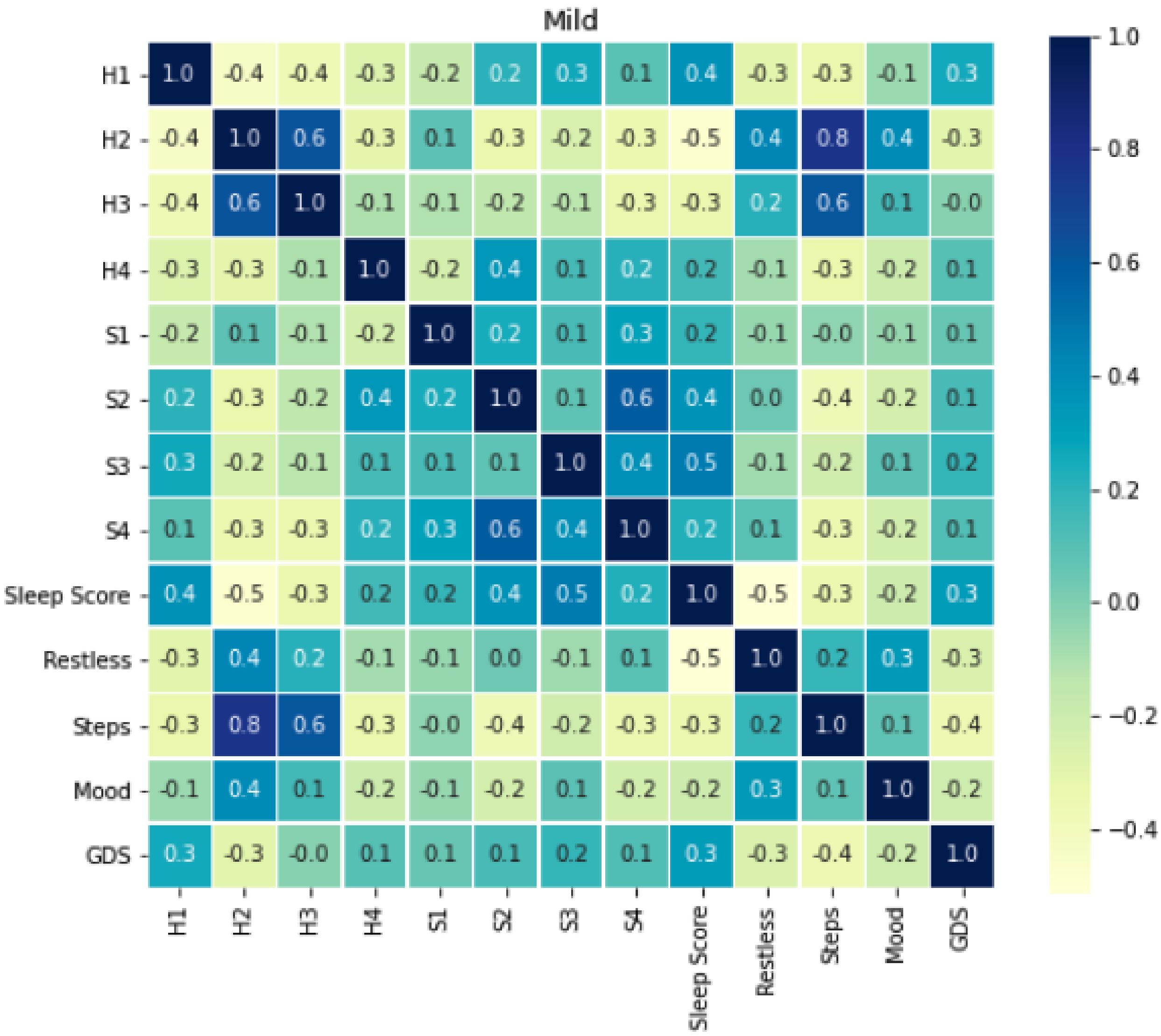



Figure 5.6: Correlations for mild depressive tendency.

Finally, a look at the heatmap for a participant with high depressive tendencies (Fig.

5.7) shows similar coefficients when compared to the participants with low and mild depressive tendencies, however with opposing polarities for most features. This links directly with Hypothesis 2, confirming that physiological behaviours for participants with higher depressive tendencies differ from those with lower depressive tendencies.

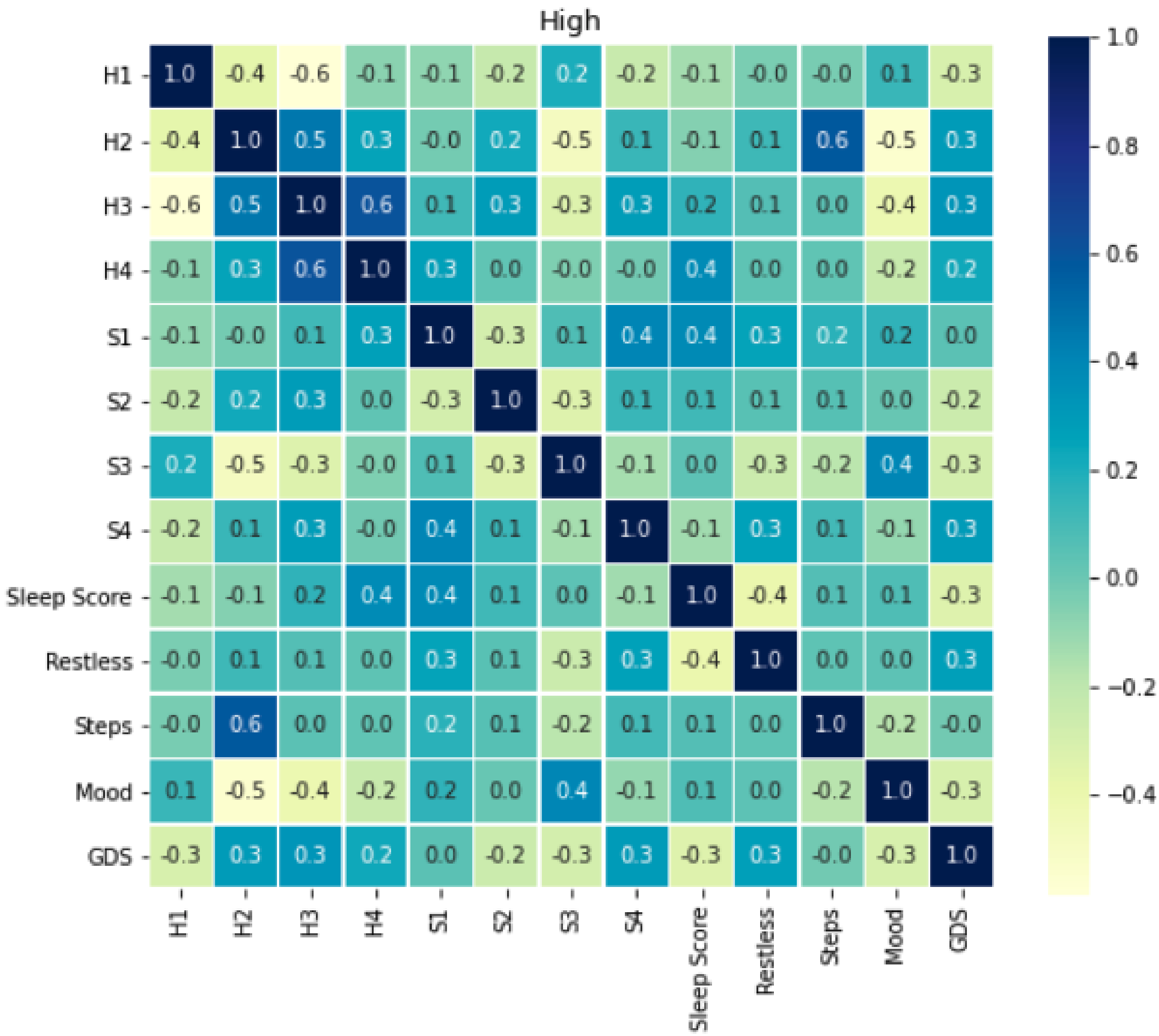



Figure 5.7: Correlations for high depressive tendency.

Figure 5.8 shows the correlation heat map for the entire dataset consisting of participant data from all 8 participants. Data for 4 excluded participants were not included in the analysis.

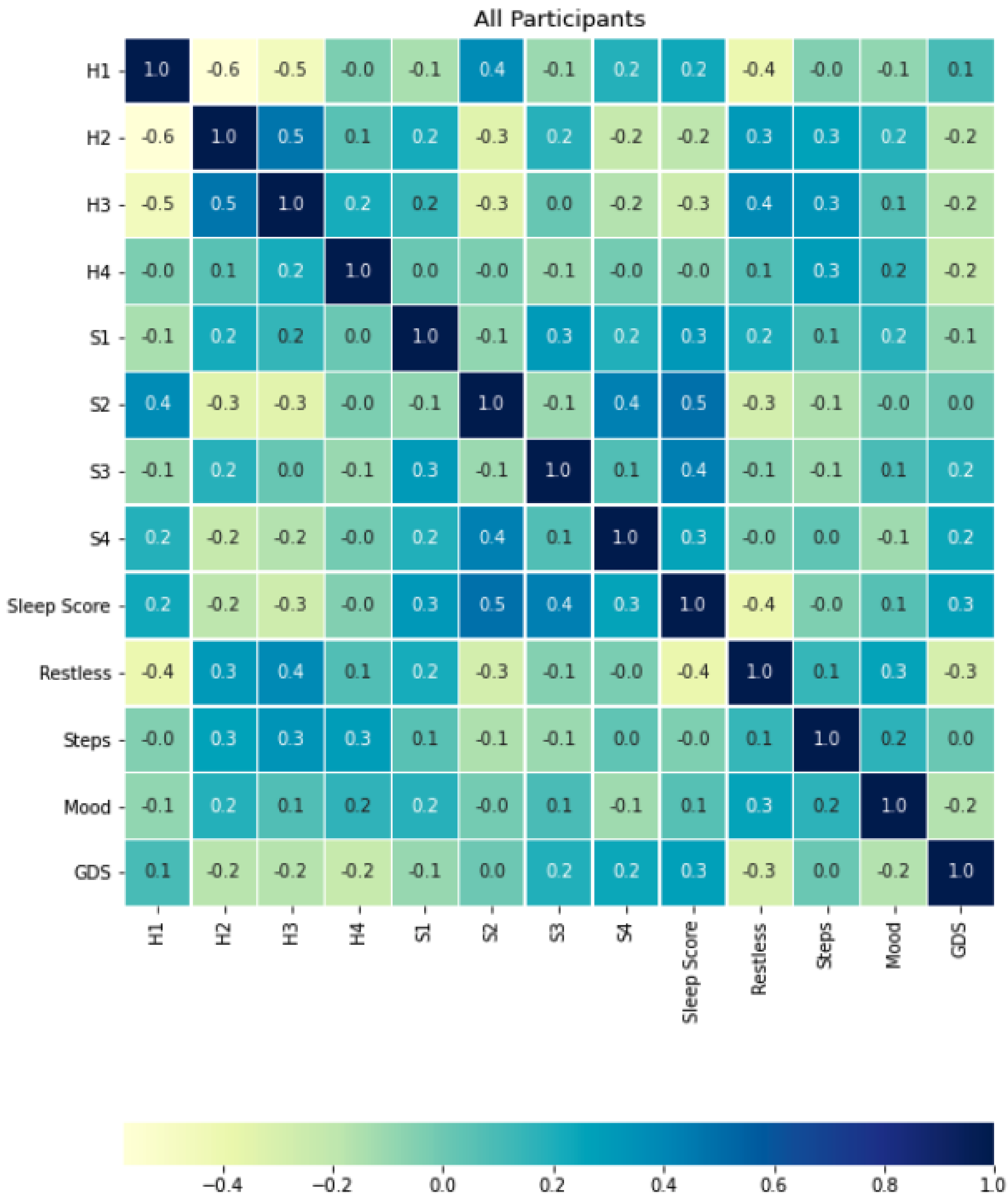



Figure 5.8: Correlations for all participant data.

The correlation plot (Fig. 5.8) for the whole dataset, correlations between physiological measures and GDS responses. To investigate whether a more balanced dataset would have a noticeable impact on the correlations, Figure 5.9 visualizes the correlation plot for a subset of the data. This subset consists of data from three participants, with one having relatively low GDS scores (0-1), one with mid-range GDS scores (3-4) and one with high GDS scores (4-9).

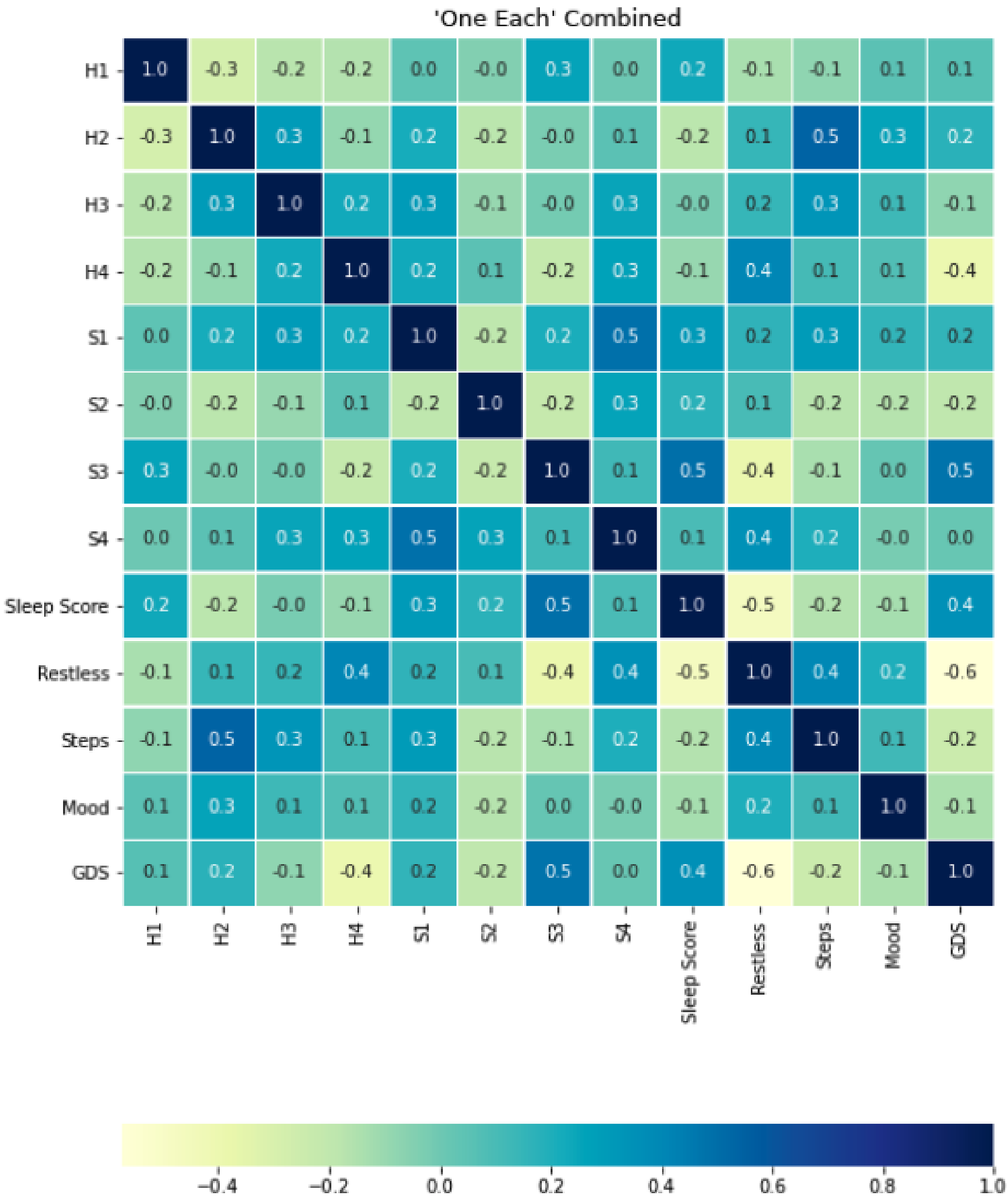



Figure 5.9: Correlations for Low-Mild-High scoring participant data.

The two heatmaps (Fig. 5.8 and 5.9) illustrate minor variations in correlations for H1, H3, Sleep Score and Mood against the GDS, while other features have slightly more noticeable coefficient value differences. These correlation plots provide visual insight and evidence that individual physiological measures or signs may not be indicative of depressive tendencies in a person. Combining these measures has the potential to provide stronger indication toward depressive tendencies in a person.

### 5.4.2 Emotion Variation

In Chapter 4, through baseline interviews and GDS scores, self-mood perception is not always accurate. Studies have also shown that verbal cues are often not indicative of inner emotions (Ekman, 2004). Further, research has also shown that patients suffering from Major Depressive Disorder (MDD), have lower Heart Rate Variations (HRV), which may mean they have lower mood fluctuations (Hartmann et al., 2019; Li et al., 2019). It has also been observed that older-aged people do not see mental health as something to talk about, or something that is treatable (CDC, 2021). As such, mood and GDS fluctuations for the 3 participants throughout the 4-week study period are shown in Figures 5.10 and 5.11.

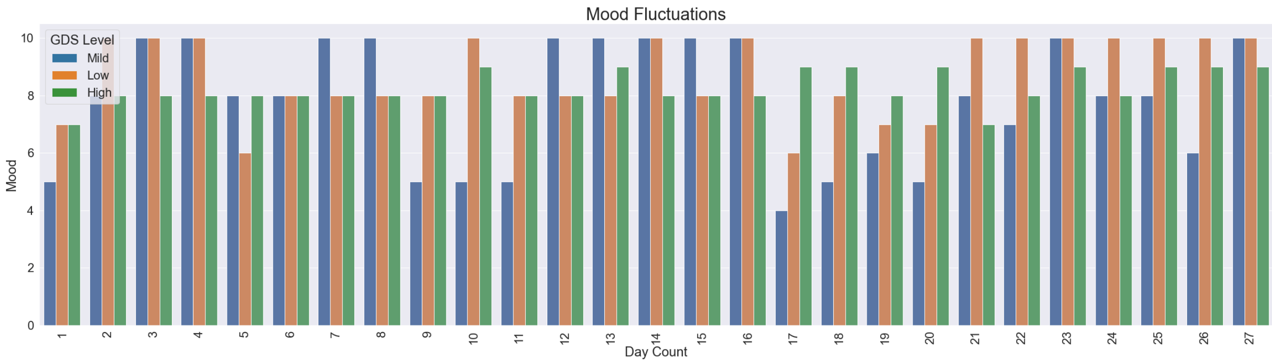



Figure 5.10: Mood fluctuations between low-mild-high scoring participants.

There are larger mood fluctuations for participant 5, with their corresponding GDS scores having a single point fluctuation. Participant 1, with relatively mid-range GDS scores, showed a similar range of mood fluctuations. On the other hand, participant 8 with the highest GDS scores showed the most stable and least fluctuating mood ratings throughout the 4-week study period. This participant also had the largest GDS score fluctuation among all 8 participants included in the post-intervention analysis. These observations highlight and strengthen previous findings that self-emotion or mood perception may not always be as indicative or reliable as presumed.

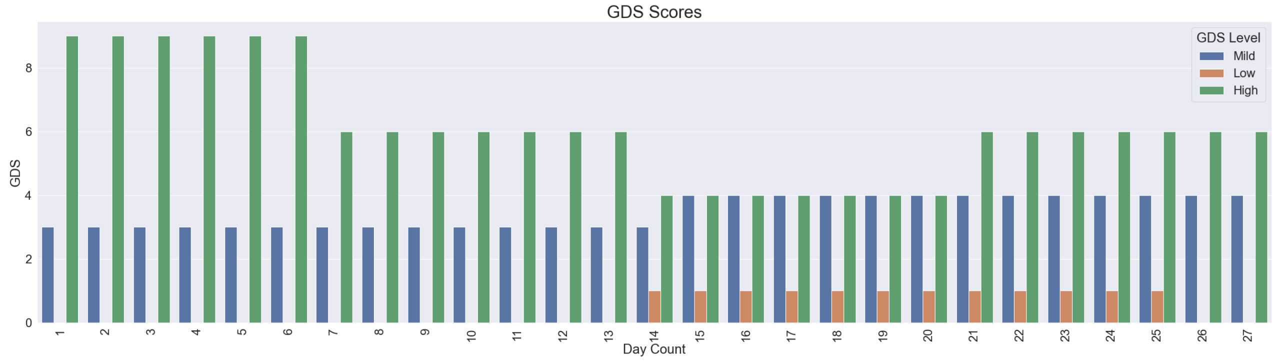



Figure 5.11: GDS score fluctuations for low-mild-high scoring participants.

Figure 5.12 additionally compares mood fluctuations between a participant with low GDS scores and a participant with high GDS scores.

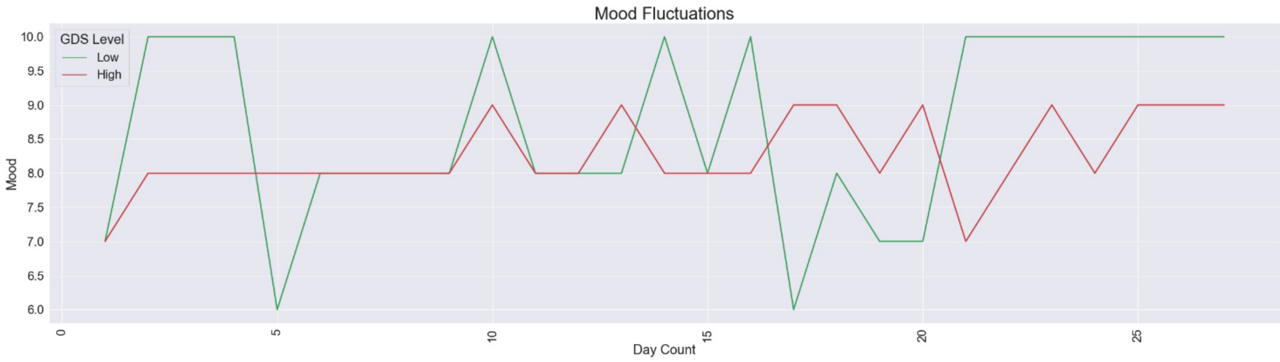



Figure 5.12: Mood fluctuations across low vs. high GDS scores.

## 5.5 Conclusion

In this chapter, a descriptive overview of physiological Fitbit data was presented to gain insight on the differences that may be reflected among participants with different levels of depressive tendency. Findings from this analysis provide the foundation for the machine learning component of the AutoMAP framework. Moving forward, using the AutoMAP dataset, a predictive regression model will be built and evaluated for its applicability in the framework as well as its performance toward depression detection. The next chapter provides insight into the predictive modelling processing.

# Chapter 6

# Predictive Modelling for Depressive Tendency Detection

This chapter provides the findings of model evaluation using the dataset presented in Chapter 5. Outcomes of the quantitative analysis provide validation on the applicability of machine learning models in the AutoMAP framework. An excerpt of the dataset used to train and test models can be viewed in Appendix C. The dataset was run through various machine learning models, particularly regression models, in order to highlight the top performing models. As the goal is to produce a non-categorical score for depressive tendency in a user, not to diagnose whether depressed or non-depressed, only regression models are evaluated in this research.

## 6.1 Introduction

In comparison to mental health applications, there is more focus toward physical health applications, such as positive activity behavioural change (Lyons et al., 2017; J. Naslund et al., 2014; Ridgers et al., 2016). Machine learning is increasingly being implemented for healthcare and has various applications such as outpatient care (Lustrek et al., 2015; NIA, 2017; Sahu et al., 2020) as well as mental health care (Garcia-Ceja et al., 2018; Kumar et al., 2021; Kyamakya et al., 2021; Randhavane et al., 2020). Kumar et al. (2021) investigated the use of ensemble machine learning methods (SVM, K-Nearest Neighbours, Decision Trees, Logistic Regression) for depression detection. The study used a ready-made dataset relating to unemployment and depression. Findings showed improved classification accuracies using ensemble methods, setting the foundations for future studies. Based on this, a different study (Randhavane et al., 2020) used video-based explicit activity recognition. They gathered affective features from gait postures and movement cues to perceive emotional states and analysed the data through Random Forest classifiers. This approach achieved a classification accuracy of 80.07%. Using contextual inputs for processing, a recent study (Choi et al., 2021) proposed emotion recognition in conversation using a convolutional neural network. Although this is an interesting direction for emotion recognition, it still requires explicit input from the user for emotion inference. Sathyanarayana et al. (2016) also explore the possibilities of using deep learning for sleep quality prediction using on raw wearable data. This study provided a foundation for possible future directions of this research. Deep learning could be implemented and compared with the outputs of a classical machine learning approach, such as random forests, decision trees and logistic regression. However, at this preliminary stage, the study is focused on using ensemble methods for depressive tendency prediction.

Research on emotion recognition using smart watches has allowed for the use various machine learning techniques to develop predictive models for emotion classification (Quiroz et al., 2018; Shu et al., 2020; Zhang et al., 2016). Data were collected under lab-based or controlled settings, requiring external emotion elicitation, along with obtrusive on-body sensors. These models mostly used heart rate or gait patterns in single or dual modalities to produce binary classifications across emotions with high performance accuracies of (1) 91.3% (neutral vs. angry), (2) 88.5% (neutral vs. happy), and (3) 88.5% (happy vs. angry), respectively; classification between happy-neutral-angry achieved an accuracy of 81.2% (Zhang et al., 2016). Our research aims to implement multivariate continuous physiological user data from a routine-based 4-week collection period with per-day summaries of user sleep, heart rate and step counts from pedometer and optical heart-rate sensors to infer depressive tendencies in a user. The implementation of the AutoMAP framework is not intended to be used as a binary, categorical or diagnostic tool. Data from these sensors has been detailed in Chapter 5. The next section presents the steps taken to develop and evaluate a predictive model for depressive tendency prediction. This model is intended to produce a continuous, non-categorical output.

## 6.2 Evaluation Setup and Measures

To determine which predictive model fits our data best, we assessed 19 models using a 10 - fold cross-validation. Shuffled data were split into 10 groups (Fig. 6.1), with 9 groups being used as training sets, and 1 group as the validation set (scikit-learn, n.d.-a). Evaluated variables included common regression metrics, namely, mean absolute error (MAE), mean square error (MSE), root mean square deviation (RMSD), R-square (R2), root mean squared logarithmic error (RMSLE), and mean absolute percentage error (MAPE). These performance measures were averaged to determine the overall performance of each evaluated model. Features were labeled with corresponding GDS and mood values to train the models for the recognition and prediction of depressive tendencies. To train, test and evaluate our predictive models, we used scikit-learn (scikit-learn, n.d.-b) and pycaret (Ali, 2020) libraries.


Figure 6.1: 10-fold cross-validation process.

### 6.2.1 Training and Selection

We compared 19 models using 10-fold cross-validation. All evaluation metrics were based on the training set, which consisted of 69.7% (150 samples) of the dataset. The remaining 30.2% (65 samples) were separated as a hold-out/test sample. Among the top 5 of the 19 evaluated models (Fig. 6.2), the gradient boosting regressor tree model outperformed other models in 4/5 metrics (MAE, MSE, RMSE, and R2). Better models have lower MAE, MSE and RSE, while having a high R2. Some R2 values are seen in the negative range, indicating that the respective model fits the data very poorly.

| | Model | MAE | MSE | RMSE | R2 | RMSLE | MAPE | TT (Sec) |
|---|---|---|---|---|---|---|---|---|
| **gbr** | Gradient Boosting Regressor | 1.0173 | 2.2774 | 1.4118 | 0.3143 | 0.5389 | 0.4662 | 0.0500 |
| **et** | Extra Trees Regressor | 1.0777 | 2.4275 | 1.4658 | 0.3095 | 0.5635 | 0.4284 | 0.1000 |
| **lightgbm** | Light Gradient Boosting Machine | 1.1122 | 2.4372 | 1.4928 | 0.2219 | 0.5718 | 0.5749 | 0.1400 |
| **rf** | Random Forest Regressor | 1.1608 | 2.7162 | 1.5570 | 0.1983 | 0.5880 | 0.4945 | 0.1340 |
| **ada** | AdaBoost Regressor | 1.3309 | 3.6117 | 1.7874 | -0.0399 | 0.6563 | 0.5103 | 0.0660 |
| **huber** | Huber Regressor | 1.3443 | 3.9783 | 1.8928 | -0.0474 | 0.6302 | 0.5586 | 0.0210 |
| **ridge** | Ridge Regression | 1.4643 | 3.6890 | 1.8499 | -0.0807 | 0.6705 | 0.6455 | 0.0100 |
| **lr** | Linear Regression | 1.4558 | 3.7220 | 1.8642 | -0.1006 | 0.6754 | 0.6863 | 0.6570 |
| **br** | Bayesian Ridge | 1.4946 | 3.8411 | 1.8907 | -0.1384 | 0.6821 | 0.6016 | 0.0120 |
| **lar** | Least Angle Regression | 1.4670 | 3.7790 | 1.8849 | -0.1495 | 0.6809 | 0.6925 | 0.0110 |
| **lasso** | Lasso Regression | 1.6480 | 4.3650 | 2.0150 | -0.2737 | 0.7308 | 0.6248 | 0.0120 |
| **en** | Elastic Net | 1.6480 | 4.3650 | 2.0150 | -0.2737 | 0.7308 | 0.6248 | 0.0160 |
| **llar** | Lasso Least Angle Regression | 1.6480 | 4.3650 | 2.0150 | -0.2737 | 0.7308 | 0.6248 | 0.0140 |
| **dummy** | Dummy Regressor | 1.6480 | 4.3650 | 2.0150 | -0.2737 | 0.7308 | 0.6248 | 0.0090 |
| **knn** | K Neighbors Regressor | 1.4320 | 3.7659 | 1.8719 | -0.2883 | 0.7016 | 0.7155 | 0.0130 |
| **omp** | Orthogonal Matching Pursuit | 1.6557 | 4.3993 | 2.0382 | -0.3588 | 0.7449 | 0.6801 | 0.0140 |
| **dt** | Decision Tree Regressor | 1.1667 | 4.1533 | 1.8824 | -0.6191 | 0.6840 | 0.6197 | 0.0120 |
| **par** | Passive Aggressive Regressor | 2.1431 | 7.0942 | 2.6034 | -1.5949 | 0.8841 | 1.2353 | 0.0120 |

Figure 6.2: Model comparison of 19 best performing regression models.

### 6.2.2 Feature Evaluation

We visualized scaled measure distributions for all users using boxplots (Fig. 5.1). A small spread indicates a similar amount of time (minutes) in the respective measure zone among all users. A more spread-out boxplot on the other hand, indicates variability in the respective zones. Outliers were not removed, to observe their impact on determining depressive tendencies.

| Features | p-values |
|---|---|
| H1 | 0.123 |
| H2 | 0.009 |
| H3 | 0.015 |
| H4 | 0.001 |
| S1 | 0.177 |
| S2 | 0.929 |
| S3 | 0.004 |
| S4 | 0.000 |
| Steps | 0.893 |
| Sleep Score | 0.000 |
| Restless | 0.000 |

Figure 6.3: Feature p-values.

We assessed the p-values for all features (Fig. 6.3). This showed statistical significance for H3, H4, S3, S4, sleep score and restlessness for predicting depressive tendencies within the sample population. These values show that certain heart rate and sleep zones can be indicative of depressive tendencies, while others may not be as relevant. P-values of 0.05 or less were deemed statistically significant. While p-values were used to reject the null hypothesis for features, they did not feature importance.

For better model interpretability, we computed feature importance across all features. To calculate the feature importance, the number of times a single feature was used as the starting point of a tree node and then divided by the total number of trees to get an average. We plotted feature importance values (Fig. 6.4) for each feature for all users. This feature importance plot shows which features contributed more to the predictive power of the model. A higher value indicates more importance over other features for prediction. A compact boxplot signifies that the respective feature was important among across all samples. Concurrently, a taller boxplot means that the data is more spread out, and signifies that the feature was only important for some samples, while not as important for others. Among all users, sleep score was the most important feature.

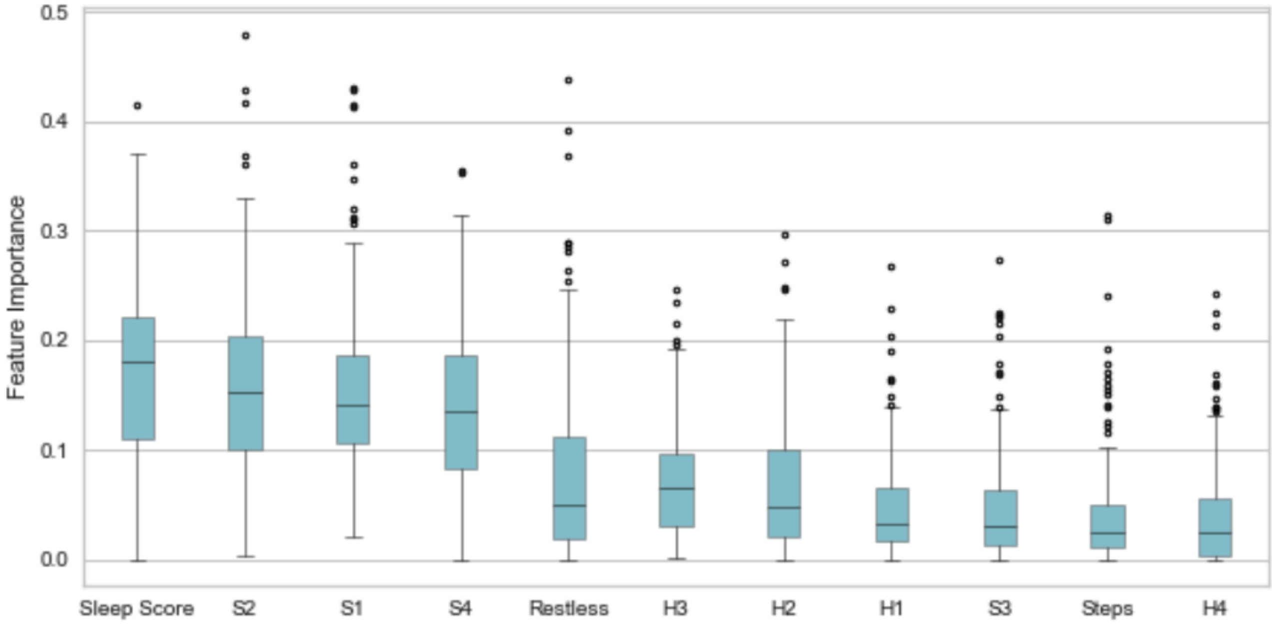



Figure 6.4: Feature importance.

## 6.3 Discussion

Our findings show that physiological signals can be used to identify depressive tendencies in older aged people using some models. We also observed activity similarities within subjects, as shown in Fig. 5.1. Although some models showed some predictive accuracy, most models did not. Given this, and the limited sample size, it is possible that significant results are a result of model overfitting rather than real predictive accuracy. As such although these preliminary results are promising, they would need to be replicated in a larger sample of older people, who exhibit a wider range of depressive tendencies.

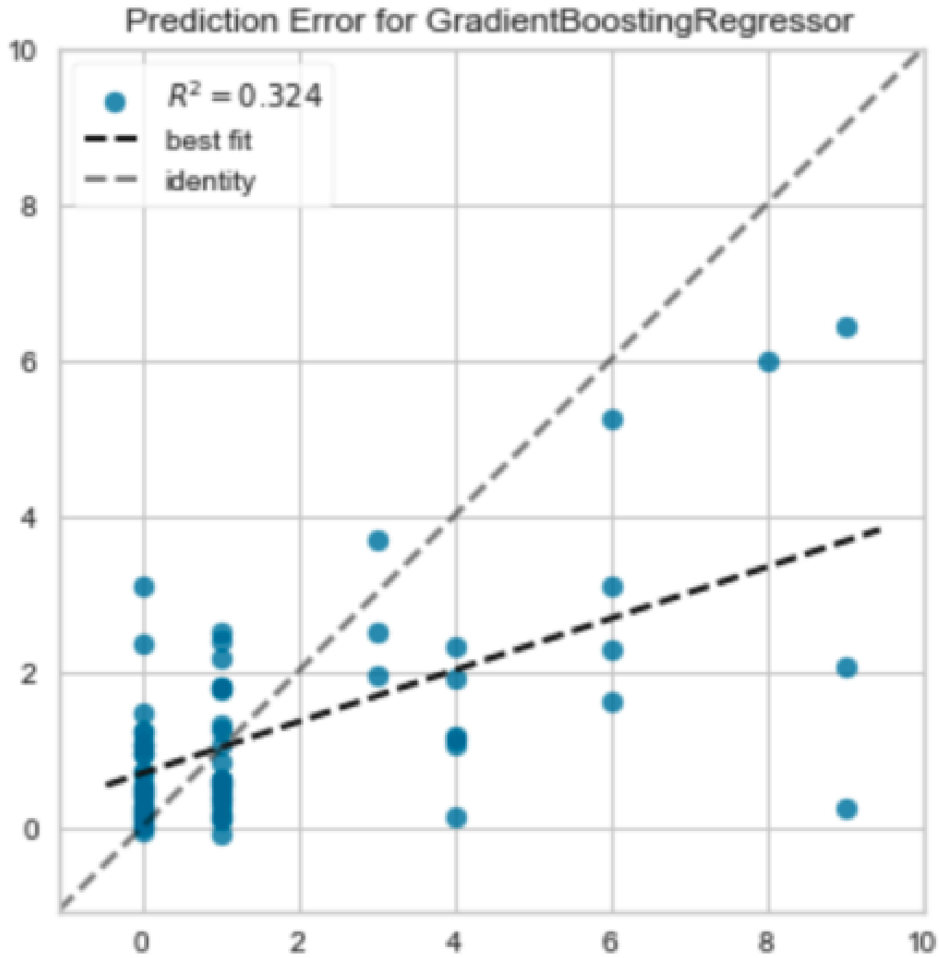



Figure 6.5: Gradient boosted regressor goodness of fit.

The distribution of heart, step count and sleep behaviors among older aged people showed similarity among movements in the sample population. A combination of heart rate zones, step count, and sleep zones together could be used to detect depressive tendencies in older aged people. Linear regression p-values (Fig. 6.3) indicate that most evaluated measures have a significant probability of identifying depressive tendencies in a user. Feature importance (Fig. 6.4) further strengthens that these measures, when combined, may be able to predict depressive tendencies in the user.

Looking at feature importance, sleep measures hold more weight in identifying depressive tendencies in older aged people. The higher the feature importance, the more predictive the feature is. Combined with step count and heart activity minutes, the ability to predict increased. Comparing model performances using features extracted from the Fitbit Alta HR smartwatch, we observed an R-square of 0.3268. On a scale of 0 to 1, the higher the R-square, the better.

Normally, R-square values ranging 0.6 - 1.0 are considered high. However, due to the nature of the dataset, a value of >0.7 in this case would indicate overfitting. In human behavior or psychology-based research, it is not uncommon for R-squared values =<0.5 as humans are unpredictable (Ballard, 2019). This is however likely to have been symptomatic of the limited size and nature of the dataset. While 0.3268 is less than ideal, it is promising given the limited dataset. It also provides a foundation for future research and direction toward the application of smart wearables in the mental health domain. Moving forward, we aim to retrain our model with a larger, more balanced dataset.

Our data collection and analysis stage came with various challenges. Three of the 12 participants did not return for the post-intervention interviews conducted upon study completion. Of the remaining 9, 2 participants had device sync issues, of which 1 participant cooperated in resyncing and the other was not responsive after continued efforts. Due to this, 4 participants and their Fitbit data were excluded from the study. As a minimum GDS score cut-off was not part of the inclusion criteria 6/8 participants had low to no depressive tendency, while the remaining 2 were on the higher end of the depressive tendency spectrum (mean score: 7/15). This led to the dataset being skewed in relatively small dataset. Despite this, we observed p-values of <0.05 for individual features, indicating an effect in predicting depressive tendencies in a person. A larger dataset may provide further insight into the statistical significance of each measure.

| | MAE | MSE | RMSE | R2 | RMSLE | MAPE |
|---|---|---|---|---|---|---|
| **0** | 0.6575 | 0.5909 | 0.7687 | 0.8506 | 0.4313 | 0.3674 |
| **1** | 1.3269 | 5.3891 | 2.3214 | 0.0346 | 0.6741 | 0.3811 |
| **2** | 1.0713 | 2.3626 | 1.5371 | 0.6990 | 0.4798 | 0.4686 |
| **3** | 0.7969 | 1.2958 | 1.1383 | 0.6060 | 0.4006 | 0.3365 |
| **4** | 0.8301 | 1.1665 | 1.0800 | 0.6334 | 0.3996 | 0.4026 |
| **5** | 1.2181 | 2.0304 | 1.4249 | 0.0752 | 0.5238 | 0.7453 |
| **6** | 1.5939 | 6.2338 | 2.4968 | 0.1720 | 0.7271 | 0.4720 |
| **7** | 0.8959 | 1.2797 | 1.1312 | -0.2305 | 0.5729 | 0.6543 |
| **8** | 0.8555 | 1.2012 | 1.0960 | 0.5641 | 0.5505 | 0.4287 |
| **9** | 0.8806 | 1.1622 | 1.0781 | -0.1369 | 0.5991 | 0.4035 |
| **Mean** | 1.0127 | 2.2712 | 1.4073 | 0.3268 | 0.5359 | 0.4660 |
| **SD** | 0.2728 | 1.8396 | 0.5393 | 0.3660 | 0.1060 | 0.1251 |

Figure 6.6: Gradient boosted regressor performance scores.

## 6.4 Conclusion

Despite challenges, preliminary analysis and model training indicate that smart wearable physiological data can potentially be used to predict depressive tendencies in users. We further hypothesize that a more balanced and larger dataset would produce stronger evidence and better validation in the direction of using smart wearables for unobtrusive mental health monitoring.

Our feasibility study, at an earlier stage found AutoMAP to be favorable to older aged people (Mughal, Raffe, Stubbs, Kneebone, et al., 2021). Older aged people noted that they would feel more comfortable knowing their mental health was being monitored, without requiring physical assistance or continual interaction with a person or device. Since it requires no explicit user interaction and was found to be primarily unobtrusive (n = 2), AutoMAP may be more impactful for those with verbal impairments or disabilities, such as intellectual disability.

CHAPTER 7

# CONCLUSIONS, LIMITATIONS AND FUTURE WORK

The goal of this research has been to develop and investigate the feasibility and acceptability of a conceptual framework that provides autonomous mental health monitoring for older aged users through minimally obtrusive data collection and little to no explicit user interaction. From this, a user study was planned, designed, and conducted for qualitative and quantitative inputs from participants. The framework was tested and assessed through (1) user device testing, (2) feasibility participant interviews, and (3) predictive model evaluation.

In Chapters 1 and 2, theoretical foundations were laid out, along with the introduction of the AutoMAP. Chapter 3 then outlined the methods and materials applied in this research. Chapter 4 presented the qualitative outcomes from a feasibility evaluation on the User Input component of the AutomMAP framework. Subsequently, Chapter 5 provided descriptive insights from exploratory analyses on quantitative physiological Fitbit data. Building on the findings from the preceding chapter, Chapter 6 explored the predictive modelling and evaluation processes that fit into the Machine Learning component of the AutoMAP framework.

This chapter delivers a high-level summary of the work presented in this thesis. The first part of this chapter summarizes the key contributions of this research. The second part of this chapter outlines the limitations encountered during the various stages of this research. Finally, the third chapter of this chapter describes the future directions of this work along with concluding statements.

## 7.1 Contributions

This thesis presented a new conceptual framework for autonomous mental health monitoring for older aged people, referred to as AutoMAP (Fig. 7.1). The AutoMAP framework was designed to incorporate and implement consumer-grade smart wearables, machine learning models and a mobile application to aid in improving the quality of life for older aged people that live alone.

Through a comprehensive literature review, strengths, gaps, and limitations of existing emotion recognition research were identified, which narrowed down and selected the least obtrusive device (the Fitbit Alta HR), while being cost-effective and easily implementable. The data collected and extracted from this device showed encouraging correlations with self-reported mood ratings and GDS scores for users. Extracted Fitbit data for all users was preprocessed and collated into a single dataset comprised of heart rate zones, sleep zones, sleep score, restlessness, step count, GDS score, and self-reported mood Likert ratings. Scores calculated from participant GDS responses were used as labels to train the machine learning model for prediction of depressive tendencies in users. A comparison of self-reported mood ratings and GDS scores showed that self-perceived emotional wellbeing may not always be as accurate and reliable. This, strengthens the need for autonomous mental health monitoring, particularly in older aged people.

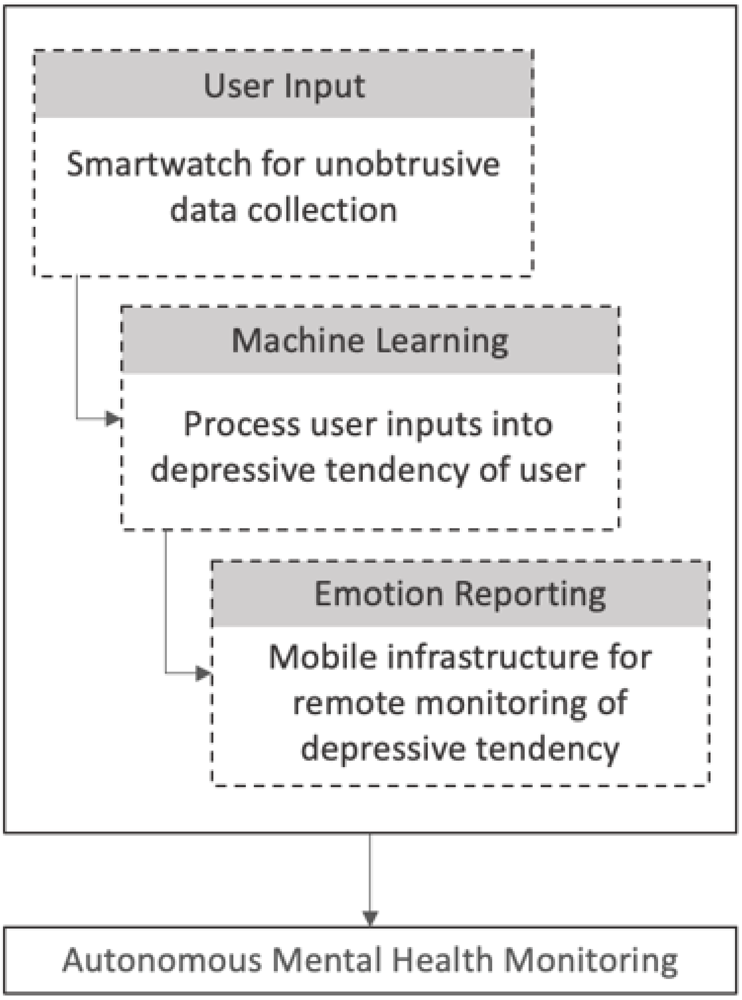



Figure 7.1: The AutoMAP framework.

Interviews with participants pre- and post- procedure were analysed suggesting that the AutoMAP framework has promising potential toward improving the quality of life of older aged people that live alone. Having their mental health monitored without explicit interaction with caregivers, or requiring constant care, would likely make them feel more comfortable and independent.

Due to the nature of GDS labels, only regression models were evaluated for our predictive model. The GDS scores range from 0 - 15, and so the AutoMAP framework was designed to process continuous values, producing a predictive score within a range aligned with the validated GDS. Additionally, as the AutoMAP framework is not intended as diagnostic tool, and binary classifications of depressed vs. non-depressed were ruled out at a very early point of this study.

The key contributions that addressed the research motivations are summarized in the next section.

## 7.2 Addressing Research Questions

This section offers an overview of how each of the research questions presented in Chapter 1 have been addressed by key contributions of this study.

### 7.2.1 Least obtrusive device for user input

The first research question was 'Will older aged people use the chosen wearable device with ease and convenience in a routine-based implementation?'

Through a comprehensive literature review in Chapter 2, a device was selected and incorporated into the AutoMAP framework. During a 4-week study period, participants wore the Fitbit Alta HR for the entire period. To assess the convenience and acceptability of this device for data collection, participants were interviewed on the feasibility and comfort of having to wear the device all day, especially when sleeping. Our analyses detailed in Chapter 4 found common themes of acceptance and convenience when using the Fitbit device. Findings from our qualitative analysis provide evidence that a less obtrusive data collection is feasible and potentially more suitable for real-world implementation and applications for emotion recognition. Inclusion of participants with diagnosed mental health illnesses or cognitive impairments may provide different inputs, which would be the ideal next step of this research.

### 7.2.2 Applicability for classifier training

The second research question was 'Can physiological data from wearable devices be indicative of depressive tendencies in users?'

Through Chapters 5 and 6, following data extraction from the Fitbit device and dataset preparation, various machine learning models were evaluated. As the Fitbit data consisted of continuous values, only regression models were evaluated as the AutoMAP framework aims to be a predictive tool, not a diagnostic one. While the models did not have ideal performance, given the limited nature of the dataset, the results are promising as a proof-of-concept validation. Additionally, the performance of the models was impacted by the imbalance in depressive tendencies of the participants, not the Fitbit data itself. It can be concluded that Fitbit or smartwatch data can potentially be used to predict depressive tendencies. Although an improved dataset would be integral in validating the full applicability and potential of Fitbit or smartwatch data being indicative of depressive tendencies is users.

### 7.2.3 Potential in reducing caregiver burden

The final research question was 'Can the proposed framework have a potential impact in reducing caregiver burden and assisting older aged persons living independently?'

While participant caregivers were not directly interviewed in this study, participants were asked about the feasibility, usability, and potential impact of the AutoMAP framework for themselves. A common theme in participants responses was that the framework would help them feel more comfortable and at ease knowing their mental health is looked after through remote monitoring. Some participants also highlighted that the framework may have a higher impact for those with disability. This was also highlighted as a gap earlier in this thesis. Post-stroke patients, non-verbal users, people with people with Autistic spectrum disorder, dementia or Alzheimer's could also benefit from the full implementation of the AutoMAP framework. This will also be explored in the future directions of this research. Through our qualitative interview analyses, it can be concluded that the user of consumer-grade smart wearables to monitor the mental health of older aged people can potentially minimize caregiver burden that may be faced by the carers of older aged people. Further investigation would be essential in addressing whether AutoMAP can effectively reduce caregiver burden.

Although this research has investigated a proof-of-concept framework and can potentially be used to autonomously monitor mental health of older aged people, the limited size of the sample population leaves room for further research toward Research Question 2. The next section will cover the challenges and limitations of this research.

## 7.3 Limitations

This research inevitably came with some challenges and limitations. As this was a pilot study, ethical constraints limited medically diverse recruitment of older aged people. The initial aim was to recruit participants with diagnosed mental health issues, Alzheimer's, and dementia. However, due to ethics constraints at this preliminary stage of the research, only generally older aged people were able to be recruited into the study. Furthermore, due to limited resources and possibly the COVID pandemic, the population sample was also fairly small. Participants were also not recruited based on their GDS scores, and as such there was an imbalance across participants with low depressive tendencies (score = 0 - 5) and mild to high depressive tendencies (score = 6 - 15).

Upon completion of the study, some participants did not return (n = 3). Due to this, their Fitbit data and pre-procedure interview data was excluded from our analyses. This impacted the model training component of the research and increased the chances of model overfitting. A more medically diverse, and a much larger sample population has the potential to produce more promising results in the future. Given the limited nature of the dataset, it would be imperative to further evaluate the reliability and capability of physiological smartwatch data in predicting depressive tendencies in users.

## 7.4 Principal Findings

The most significant findings of this thesis include the development of a conceptual framework that delivers autonomous mental health monitoring for older aged people, a comprehensive literature review and the validation of using consumer grade smart wearables as a means of unobtrusive data collection. The main components of this framework were tested and assessed through participant feedback and evaluation of predictive performance of the machine learning model.

A comprehensive systematic review presented the latest trends in mental health, particularly depression, emotion recognition using smart devices, types of sensors used, common settings for data collection and emotion elicitation. Mixed design analyses led to the validation of consumer grade smart wearables being acceptable and feasible for use as a means of unobtrusive data collection on the use of depressive tendency monitoring among older aged people. Finally, incorporating smart devices and applied machine learning to aid in potentially alleviating mental health episodes in the older aged generation by means of monitoring and early detection, a conceptual framework was designed and evaluated.

As introduced in Chapter 1, this research sought out to answer three primary questions:

- Will older aged people use the chose wearable device with ease and convenience in a routine-based implementation?
- Can physiological data from wearable devices be indicative of depressive tendencies in users?

- Can the proposed framework have a potential impact in reducing caregiver burden and assisting older aged persons living independently?

While addressing these questions, a series of experiments and analyses were conducted. These experiments and analyses were designed to systematically build upon each other.

(1) Initially, a thorough literature review was conducted. This review presented a comparative overview of the current state of research in the domain of smart wearables for health. The primary purpose of this thesis was to extend and contribute to the body of knowledge on smart devices, digital health, and ageing. The AutoMAP framework was also developed through this review, which served as a foundation for the remainder of the research. Findings from this systematic literature review and the AutoMAP framework were presented at SeGAH '20 as a conference paper (Mughal et al., 2020).

(2) Through a 4-week exploratory closed cohort study, a feasibility evaluation was administered with 12 older aged people using Fitbit Alta HR smart watches. This investigation determined the usability, feasibility, and applicability of using smart wearables, particularly Fitbit Alta HR smartwatches for unobtrusive, implicit user interaction. Qualitative findings from this evaluation have been submitted to the Journal of Medical Internet Research, and is currently in the final stages of review (Mughal, Raffe, Stubbs, Kneebone, et al., 2021).

(3) Following the 4-week study, physiological data for all participants were extracted and investigated for correlations between depressive tendencies and physiological behaviours. Data from the Fitbit were used to train and evaluate machine learning models for depressive tendency prediction. Calculated depressive tendency scores were used as data labels for supervised learning. Quantitative findings and descrip-

tive insights for the resultant dataset have been accepted as a conference paper at SeGAH '22. These findings provide evidence that physiological behaviours can be indicative of depressive tendencies users and highlighted the applications of machine learning models for depressive tendency detection through Fitbit data.

## 7.5 The AutoMAP Framework

Our aim through this research was to use Fitbit smartwatches, machine learning models and a mobile reporting application to aid in reducing caregiver burden by monitoring user mental health remotely. As such, we proposed AutoMAP (Fig. 1.3). With this framework, we used smart watches for data collection (Fitbit Alta HR). This was an unobtrusive method to collect data and requires minimal to no explicit user interaction. Extracted physiological data (sleep patterns, steps, and heart rate patterns) was then processed and run through a predictive model to determine depressive tendencies in the user. Calculated scores would then be reported to an allocated caregiver via a mobile app. To address privacy concerns, the user would be able to decide the depth of information shared with their caregiver, with the depressive score being the bare minimum. Our feasibility validation of the AutoMAP framework achieved an 87.5% retention and 75% acceptability rate. The protocols followed in this study are detailed in Chapter 4.

## 7.6 Future Work

The work presented in this thesis provides several potential directions for future work. For the user input component of the framework, in terms of devices for data collection, different consumer-grade smart wearables can be compared and evaluated. In this research, based on comprehensive literature reviews and cost comparison, the Fitbit Alta HR was selected. There are many competing smart watch brands that may or may not provide more accurate user physiological measurements. Branching out from just smart watches, other internet of things (IoT) devices and smart home technologies could also be evaluated for potential applicability to the AutoMAP framework. As discussed in Chapter 2, wireless internet routers have can recognize activities through walls, this paves the way for assessment of emotion and possibly mental health.

A larger and medically diverse participant cohort could be included in the study to investigate the impact of the framework. In particular, participants with diagnosed depression, dementia, Alzheimer's, autism spectrum disorder or post-stroke patients could provide more insight on the accuracy and usefulness of the AutoMAP framework. More participants, spread out more evenly along the depressive tendency (GDS) spectrum could improve model performance in predicting depressive tendencies in a person. Further, this research consisted of more female participants (n = 8) as opposed to male (n = 4). Four female participants were excluded from the study, resulting in a gender balance during further analysis. Gender impact on depression, reduction of caregiver burden and the performance of the AutoMAP framework could be investigated in more depth with a participant pool consisting of an equal number or men and women. Other factors to potentially study could include (1) socioeconomic status, (2) cultural background, and (3) geographical location. These factors may increase or reduce the usability, feasibility, acceptability, and overall impact of the AutoMAP framework.

Data collection for this pilot study was conducted over a 4-week period. This duration could be extended for further framework validation and feasibility testing. Extended user case studies with older aged people and their respective caregivers would allow for a more in-depth analysis on how effectively the AutoMAP framework can aid in reducing caregiver burden. Future replications of this study should aim to branch out, not only to a more medically diverse participant population, but also investigate the potential applicability for detecting depressive tendencies in users of any age. Other sensor data could also be investigated for their effect on the performance accuracy in detecting depressive tendencies in users, such as gyroscope data, accelerometer data, and stride patterns.

The AutoMAP framework presented in this thesis is a step toward autonomous mental health monitoring through least obtrusive means. The framework incorporates cost-effective, easily implementable, and convenient smartwatches with machine learning and mobile applications to increase comfort and the quality of life when living alone, while potentially alleviating caregiver burden. Even in the face of multiple obstacles and challenges during various phases of this research, this research lays out a foundation toward autonomous mental health monitoring among older-aged people.

# Appendix A

# APPENDIX

UNIVERSITY OF TECHNOLOGY SYDNEY

**PARTICIPANT INFORMATION SHEET**

***Autonomous Monitoring System for Older Aged Persons (HREC Ref ETH20-4912)***

WHO IS DOING THE RESEARCH?
Miss Fiza Mughal (PhD Candidate at UTS)
Dr. Jaime Garcia (Expert in Serious Games for Health)
Dr. William Raffe (Expert in Games and Artificial Intelligence)
Dr. Peter Stubbs (Expert in Physiotherapy)
Prof. Ian Kneebone (Expert in Clinical Psychology)

WHAT IS THIS RESEARCH ABOUT?
This research is about tracking the everyday movements (sleep, walk, and heart rate) of older aged people, and use these to develop a computer program to automatically monitor well-being.

This is the first stage of the study, in which we will use a Fitbit (a watch-like band that measures your physical activity) to collect daily data and develop a program that can use this data create a tool that automatically detects changes in emotions.

The automatic monitoring of mood and can inform the families of older aged persons when they may require further help. The proposed final product may provide peace of mind to your families knowing that your wellbeing is being monitored while also allowing you to live more independently.

WHY HAVE I BEEN ASKED?
You have been invited to participate in this study because you are 65 years old (or more) and living alone. You are able to provide daily information about your activities, well-being and diet so that we are able to develop an automated program that can monitor independent older aged people's mental health using everyday movements such as walking and sleeping.

You have also been asked because you:
1. Have a computer, mobile phone, and internet access
2. Have family members concerned about your wellbeing

IF I SAY YES, WHAT WILL IT INVOLVE?
If you decide to participate, I will invite you to:
1. Wear a Fitbit smartwatch for a duration of 4 weeks.
2. Participate in an interview before the 4-week study.
3. Answer questions online every day. These questions should take about 5 to 10 minutes. For these questions, you will be required to provide input on some parts of your daily life like:
    a. Food preferences
    b. Activities performed
    c. How your mood was through the day
4. Fill a brief questionnaire once per week. This survey should take up to 10 minutes.
5. After the 4-week period, participate in a short interview on your experiences during the 4-week period and your opinions on possible improvements.

Over the 4-week period, we will require a maximum of 6 hours of your undivided time and attention.
Prior to and following the 4-week period you will be interviewed on your daily life, how you are feeling and your regular activities.

ARE THERE ANY RISKS/INCONVENIENCE?
Yes, there are some risks/inconvenience. You might be slightly inconvenienced by answering some questions every day or having to wear the Fitbit device over the 4-week period. You may experience slight emotional distress in the interviews and filling out the questions every day. If this occurs, we encourage you

*Participant Information and Consent Form Version 3, 4th Nov 2020*

*Page 1 of 3*

Figure A.1: Participant Information: Page 1

UTS UNIVERSITY OF TECHNOLOGY SYDNEY

to inform Miss Fiza Mughal and see a GP for further advice. Please note that any further assessment referred by your GP may be covered by Medicare.
You are free to opt out of the research at any stage. You will also be provided with a list of contacts and resources you can use in the event of any distress.

### DO I HAVE TO SAY YES?

Participation in this study is voluntary. It is completely up to you whether or not you decide to take part.

### WHAT WILL HAPPEN IF I SAY NO?

If you decide not to participate, it will not affect your relationship with the researchers or the University of Technology Sydney. If you wish to withdraw from the study once it has started, you can do so at any time without having to give a reason, by contacting Miss Fiza Mughal [redacted] or [redacted]

If you withdraw from the study, you will just need to send the device back to us. We will thank you for your time and we will not contact you again.

If you decide to leave the research project, we will not collect additional personal information from you, although personal information already collected will be retained to ensure that the results of the research project can be measured properly and to comply with law. You should be aware that data collected up to the time you withdraw will form part of the research project results. If you do not want them to do this, you must tell them before you join the research project.

### CONFIDENTIALITY

By signing the consent form, you consent to the research team collecting and using personal information about you for the research project. All this information will be treated confidentially. Your contact information will be kept in a separate database only for following up with you in case of any distress and to send you the device upon study commencement. Your activity data will be identified only with a participant number, which be issued to you upon commencement. Your information will only be used for the purpose of this research project.

We would like to store your information for future use in research projects that are an extension of this research project. In all instances your information will be treated confidentially.

We plan to publish the results in academic and medical journals. In any publication, information will be provided in such a way that you cannot be identified.

### WHAT IF I HAVE CONCERNS OR A COMPLAINT?

If you have concerns about the research that you think I or my supervisor(s) can help you with, please feel free to contact on:

Miss Fiza Mughal [redacted]
Dr Jaime Garcia [redacted]
Dr. William Raffe [redacted])
Dr. Peter Stubbs [redacted]
Prof. Ian Kneebone [redacted]

You will be given a copy of this form to keep.

**NOTE:**
This study has been approved in line with the University of Technology Sydney Human Research Ethics Committee [UTS HREC] guidelines. If you have any concerns or complaints about any aspect of the conduct of this research, please contact the Ethics Secretariat on ph.: +61 2 9514 2478 or email: Research.Ethics@uts.edu.au] and quote the UTS HREC reference number. Any matter raised will be treated confidentially, investigated and you will be informed of the outcome.

*Participant Information and Consent Form Version 3, 4th Nov 2020*

*Page 2 of 3*

Figure A.2: Participant Information: Page 2

UTS UNIVERSITY OF TECHNOLOGY SYDNEY

***CONSENT FORM***

**Autonomous Monitoring System for Older Aged Persons (HREC Ref ETH20-4912)**

I ____________________ *[participant's name]* agree to participate in the research project Autonomous Monitoring System for Older Aged Persons (HREC Ref ETH20-4912) being conducted by Fiza Mughal [redacted].

I have read the Participant Information Sheet, or someone has read it to me in a language that I understand.

I understand the purposes, procedures and risks of the research as described in the Participant Information Sheet.

I understand I may also contact my usual healthcare provider, such as my doctor, if I become distressed.

I have had an opportunity to ask questions and I am satisfied with the answers I have received.

I freely agree to participate in this research project as described and understand that I am free to withdraw at any time without affecting my relationship with the researchers or the University of Technology Sydney.

I understand that I will be given a signed copy of this document to keep.

I agree that the research data gathered from this project may be published in a form that:

☐ May be used for future research purposes

I am aware that I can contact Miss Fiza Mughal if I have any concerns about the research.

________________________________________ ____/____/____

Name and Signature [participant] Date

________________________________________ ____/____/____

Name and Signature [researcher or delegate] Date

*Participant Information and Consent Form Version 3, 4th Nov 2020*

*Page 3 of 3*

Figure A.3: Participant Information: Page 3

# Checklist

**First Meeting:**

- Introductory discussion
- Briefing
- Device usage and maintenance training
- Walkthrough document pack:
    - Fitbit Easy Guide
    - Emotional Distress Contact List
- Schedule Reminder preference
    - SMS Reminder
    - Self-alarms
- Schedule tentative completion date

Tentative Completion Date: ___________________

**Tracking Chart**

<table>
<tr><td>Daily Questions</td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td></tr>
<tr><td>Weekly Questions</td><td colspan="7"></td><td colspan="7"></td><td colspan="7"></td><td colspan="7"></td></tr>
</table>

**Final Meeting:**

- Closing discussion
- Feedback

**Remember:**

- You can take the Fitbit off when showering and for an hour or two (if you feel necessary) every day
- Fill out the daily questionnaire at the same time daily
- Fill out the weekly questionnaire at the same day of every week
- Keep your Fitbit on when you sleep
- Keep your Fitbit charged
- Keep your Fitbit dry

Figure A.4: Reminder checklist

Fitbit Easy Care Guide

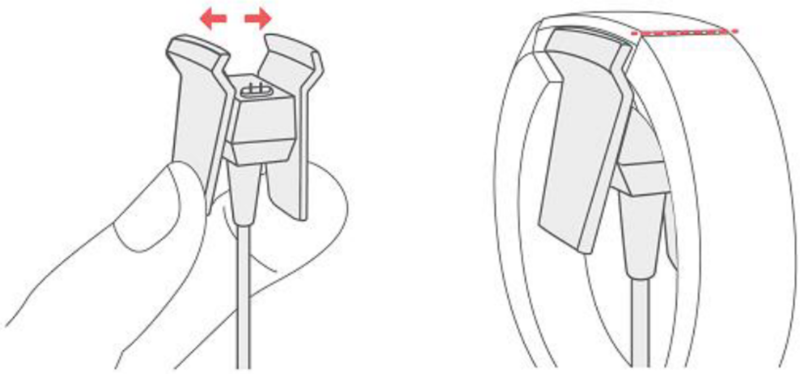
fitbit

Wear & Care

## Charge Alta HR

To charge your tracker:

1. Plug the charging cable into the USB port on your computer or a UL-certified USB wall charger.
2. Clip the other end of the charging cable to the port on the back of the tracker. The pins on the charging cable must lock securely with the port. You'll know the connection is secure when the tracker vibrates and you see a battery icon on your tracker's display. The battery icon disappears after three seconds.



Charging fully takes about one to two hours. While the tracker charges, you can tap it to check the battery level. A fully charged tracker shows a solid battery icon.

Figure A.5: Fitbit Easy Care Guide: Page 1

Fitbit Easy Care Guide

## Keep It Dry

- While Fitbit devices are water resistant*, it's not good for your skin to wear a wet band for long periods of time.
- If your elastomer band gets wet—like after sweating or showering—rinse and dry it thoroughly before putting it back on your wrist.
- Be sure your skin is dry before you put your band back on.
- Fitbit products are made to be water resistant. Fitbit Ace 2, Charge 3, Flex 2, Ionic, Inspire family and Versa family can be worn in the lake, the pool or the ocean, and Fitbit Ace can be worn in the shower. Still, it's important to thoroughly dry your device and elastomer band and remove any debris after wear in the water to avoid skin irritation. Please note that non-elastomer accessories should not be worn in water.
- To minimise damage to your tracker or watch, avoid any direct contact with sunscreen or insect repellent sprays. Remove your device while applying these sprays.

## Don't Wear It Too Tight

- Make sure your band isn't too tight. Wear the band loosely enough that it can move back and forth on your wrist.
- If you use any of our Fitbit trackers or watches with HR tracking, you can get better heart rate readings during exercise by wearing the band so it's secure, but not too tight, and wearing it higher on your wrist (about 2-3 finger widths above your wrist bone). Lower the band on your wrist and loosen it after exercise. Learn more.

## Give Your Wrist a Rest

- Prolonged rubbing and pressure may irritate the skin, so give your wrist a break by removing the band for an hour after extended wear.

Figure A.6: Fitbit Easy Care Guide: Page 2

Fitbit Easy Care Guide

## CAUTION!

If you feel moisture on your wrist, take off the device and let it the area dry before putting the device back on.

If you feel any discomfort or itchiness on your wrist, take off the device and inform us.

If you feel a rash coming on, remove the device from your wrist and see a GP for further advice.

The above risks are rare and unlikely. However, it is important to follow the guidelines given to you in order to avoid any of the above issues.

It is very important that you:

1. Keep the device dry
2. Do not wear the device too tight
3. Remove the device for a short time once a week
4. Do not wear the device when you go for a shower or a bath

You are welcome to contact us at any point of this study if you need more information, would like to inform us of anything or if you have any concerns about this research.

Miss Fiza Mughal [redacted]

Dr. Jaime Garcia [redacted]

Dr. William Raffe [redacted]

Dr. Peter Stubbs [redacted]

Prof. Ian Kneebone [redacted]

If you do not wish to proceed further with the study, please inform us. You will need to simply send the device back, for which you have been provided with a prepaid return envelope. We will thank you and not contact you further.

Figure A.7: Fitbit Easy Care Guide: Page 3

UTS

## Emotional Distress Resource List

We care about your wellbeing and understand that commenting and talking about / monitoring your mood and everyday wellbeing can be difficult. If this is the case, please communicate with Miss Fiza Mughal [redacted] if you are feeling overwhelmed or if there is something you would like to chat about.

Below are some other options that can be utilized.

Services/resources that are available:

- NSW Elder Abuse Helpline (1800 628 221)
    - https://www.ageingdisabilitycommission.nsw.gov.au/
- Older Australians COVID-19 Support Line (1800 171 866)
- Mental Health Line (1800 011 511)
- Mental Health Online (https://www.mentalhealthonline.org.au/)
- 24/7 Crisis Lines
    - Emergency Triple 000
    - Beyond Blue (1300 22 46 36)
    - Lifeline Australia (131 114 or https://www.lifeline.org.au/)
- Mindspot (https://mindspot.org.au/)
    - contact@mindspot.org.au
    - 1800 61 44 34
- myCompass (https://www.mycompass.org.au/)
- SANE Australia (https://www.sane.org/services/help-centre)
    - helpline@sane.org
    - 1800 187 263
- WayAhead (https://directory.wayahead.org.au/)
    - info@wayahead.org.au
    - (02) 9339 6000
- UTS Psychology Clinic (https://www.uts.edu.au/about/graduate-school-health/clinical-psychology/what-we-do/uts-psychology-clinic)

**Detailed Resources Accessible at:**

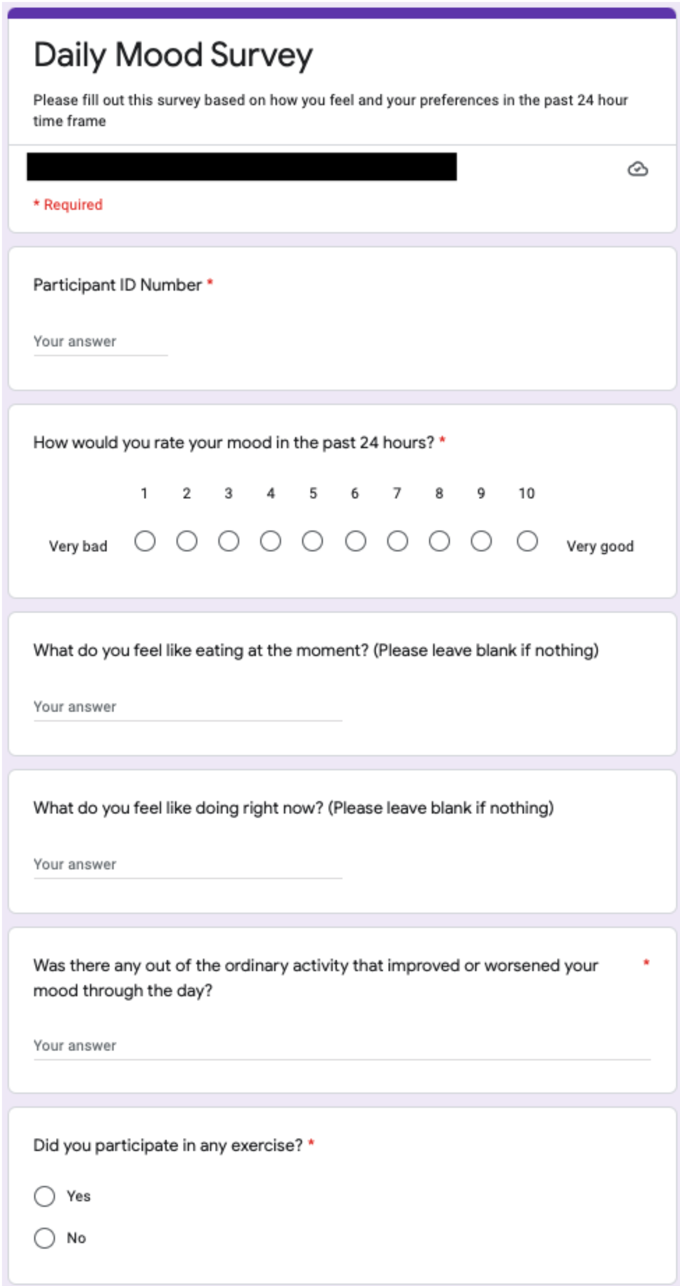
https://www.health.nsw.gov.au/mentalhealth/services/Pages/support-contact-list.aspx

https://www.health.nsw.gov.au/mentalhealth/Pages/Mental-Health-Line.aspx

Figure A.8: Emotional Distress Resource List

# Appendix B

# APPENDIX



Daily Mood Survey

Please fill out this survey based on how you feel and your preferences in the past 24 hour time frame

[redacted]

* Required

Participant ID Number *

Your answer

How would you rate your mood in the past 24 hours? *

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Very bad | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | Very good |

What do you feel like eating at the moment? (Please leave blank if nothing)

Your answer

What do you feel like doing right now? (Please leave blank if nothing)

Your answer

Was there any out of the ordinary activity that improved or worsened your mood through the day? *

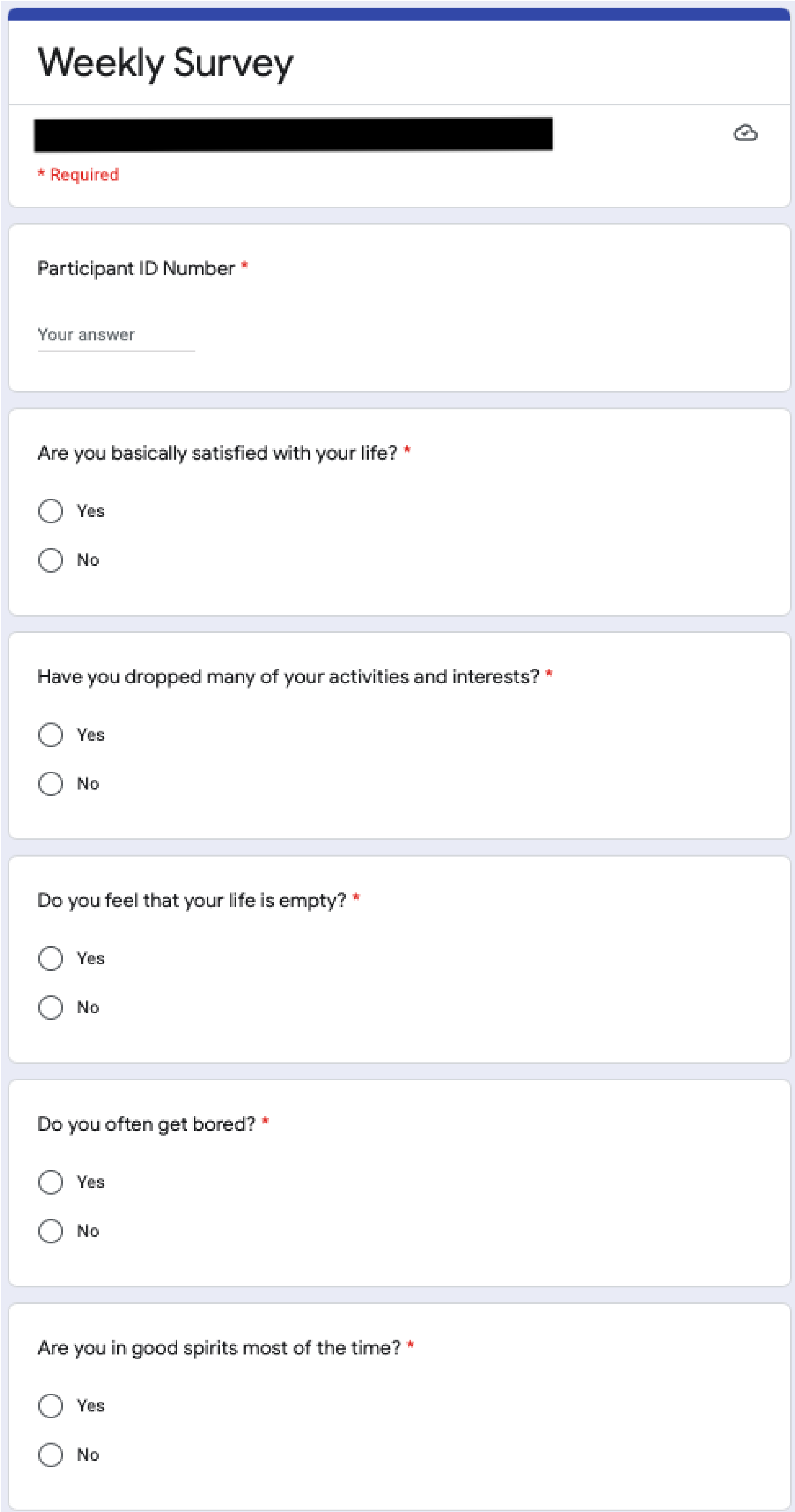
Your answer

Did you participate in any exercise? *

○ Yes

○ No

Figure B.1: Online self-reported mood Likert Questionnaire



# Weekly Survey

[redacted]

* Required

Participant ID Number *

Your answer

Are you basically satisfied with your life? *

- Yes
- No

Have you dropped many of your activities and interests? *

- Yes
- No

Do you feel that your life is empty? *

- Yes
- No

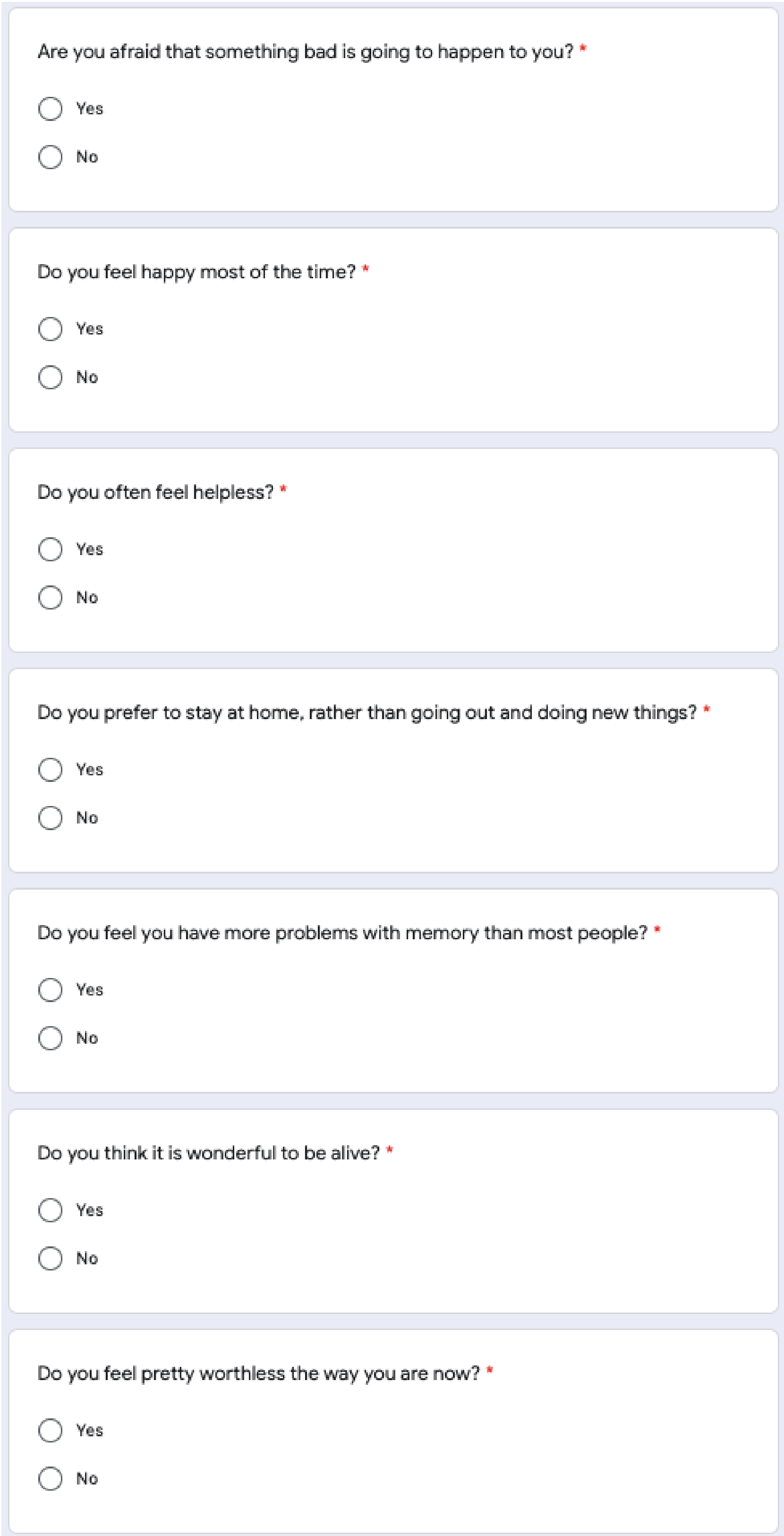
Do you often get bored? *

- Yes
- No

Are you in good spirits most of the time? *

- Yes
- No

Figure B.2: Online GDS Questionnaire - Part 1



Are you afraid that something bad is going to happen to you? *

Yes

No

Do you feel happy most of the time? *

Yes

No

Do you often feel helpless? *

Yes

No

Do you prefer to stay at home, rather than going out and doing new things? *

Yes

No

Do you feel you have more problems with memory than most people? *

Yes

No

Do you think it is wonderful to be alive? *

Yes

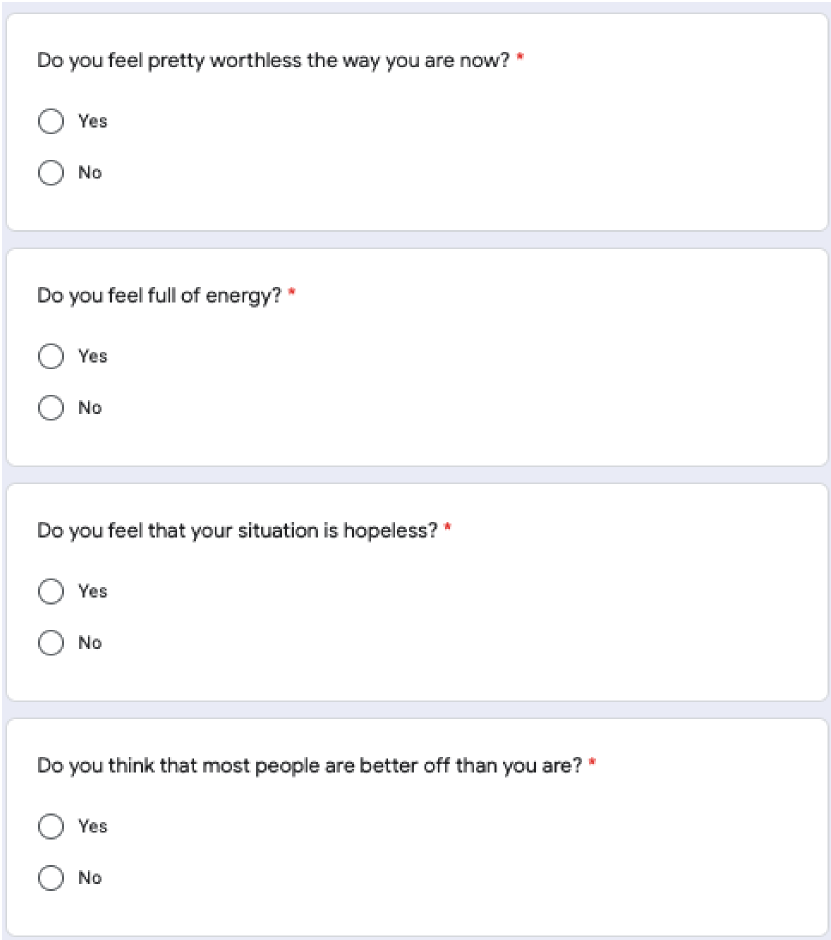
No

Do you feel pretty worthless the way you are now? *

Yes

No

Figure B.3: Online GDS Questionnaire - Part 2



Do you feel pretty worthless the way you are now? *

Yes

No

Do you feel full of energy? *

Yes

No

Do you feel that your situation is hopeless? *

Yes

No

Do you think that most people are better off than you are? *

Yes

No

Figure B.4: Online GDS Questionnaire - Part 3

# Appendix C

# Appendix

| Date | H1 | H2 | H3 | H4 | S1 | S2 | S3 | S4 | Sleep Score | Restless | Steps | Mood | GDS |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 12/2/20 | 566 | 445 | 39 | 0 | 26 | 263 | 42 | 37 | 72 | 0.06919946 | 14110 | 5 | 3 |
| 12/3/20 | 527 | 281 | 2 | 0 | 11 | 124 | 28 | 36 | 72 | 0.06919946 | 9043 | 8 | 3 |
| 12/4/20 | 549 | 658 | 28 | 0 | 47 | 293 | 67 | 54 | 61 | 0.12030075 | 8486 | 10 | 3 |
| 12/5/20 | 1080 | 130 | 1 | 0 | 29 | 350 | 79 | 60 | 83 | 0.06763285 | 6004 | 10 | 3 |
| 12/6/20 | 794 | 473 | 5 | 0 | 45 | 312 | 68 | 51 | 88 | 0.08289612 | 9457 | 8 | 3 |
| 12/7/20 | 769 | 440 | 17 | 0 | 39 | 283 | 28 | 60 | 65 | 0.06388527 | 12209 | 8 | 3 |
| 12/8/20 | 632 | 611 | 1 | 0 | 41 | 261 | 92 | 68 | 81 | 0.07649667 | 13838 | 10 | 3 |
| 12/9/20 | 342 | 427 | 24 | 0 | 35 | 280 | 59 | 53 | 75 | 0.08006852 | 7862 | 10 | 3 |
| 12/10/20 | 462 | 286 | 10 | 0 | 36 | 278 | 63 | 54 | 76 | 0.08188003 | 14367 | 5 | 3 |
| 12/11/20 | 966 | 260 | 18 | 0 | 48 | 371 | 96 | 65 | 85 | 0.06875544 | 11707 | 5 | 3 |
| 12/12/20 | 667 | 259 | 23 | 3 | 23 | 304 | 109 | 64 | 83 | 0.0735146 | 8635 | 5 | 3 |
| 12/13/2020 | 870 | 341 | 5 | 0 | 16 | 242 | 94 | 44 | 75 | 0.08491762 | 6496 | 10 | 3 |
| 12/14/2020 | 994 | 91 | 0 | 0 | 28 | 280 | 100 | 47 | 81 | 0.07683864 | 4438 | 10 | 3 |
| 12/15/2020 | 923 | 188 | 4 | 0 | 28 | 333 | 64 | 59 | 84 | 0.06638567 | 5044 | 10 | 3 |
| 12/16/2020 | 718 | 532 | 29 | 0 | 25 | 240 | 88 | 44 | 78 | 0.0721519 | 14590 | 10 | 4 |
| 12/17/2020 | 581 | 605 | 70 | 0 | 14 | 258 | 67 | 41 | 77 | 0.08525034 | 15040 | 10 | 4 |
| 12/18/2020 | 1001 | 313 | 11 | 0 | 32 | 283 | 74 | 53 | 85 | 0.0565046 | 9720 | 4 | 4 |
| 12/19/2020 | 984 | 275 | 12 | 0 | 42 | 299 | 35 | 34 | 82 | 0.0501836 | 7316 | 5 | 4 |
| 12/20/2020 | 906 | 283 | 10 | 0 | 12 | 304 | 82 | 56 | 78 | 0.08250825 | 7497 | 6 | 4 |

Figure C.1: Data excerpt for training dataset

# Bibliography

Adib, F. (2019, May 30). Household Radar Can See Through Walls and Knows How You're Feeling. IEEE Spectrum. https://spectrum.ieee.org/household-radar-can-see-through-walls-and-knows-how-youre-feeling

Ahmad, F. H., Kadir, R. S. S. A., Murat, Z. Hj., Johari, N. S. (2020). The investigation of electromagnetic radiation (EMR) for normal exercise and intense exercise difference in gender. AIP Conference Proceedings, 2306(1), 020038. https://doi.org/10.1063/5.0035056

Ali, M. (2020). PyCaret. PyCaret. https://pycaret.org/

Ali, M., Machot, F. A., Mosa, A. H., Jdeed, M., Machot, E. A., Kyamakya, K. (2018). A Globally Generalized Emotion Recognition System Involving Different Physiological Signals. Sensors (Basel, Switzerland), 18(6), E1905. https://doi.org/10.3390/s18061905

Bai, Y., Tompkins, C., Gell, N., Dione, D., Zhang, T., Byun, W. (2021). Comprehensive comparison of Apple Watch and Fitbit monitors in a free-living setting. PLOS ONE, 16(5), e0251975. https://doi.org/10.1371/journal.pone.0251975

Ballard, C. (2019). An Ode To R-Squared. https://towardsdatascience.com/an-ode-to-r-squared-804d8d0ed22c

Barrett, L. F., Adolphs, R., Marsella, S., Martinez, A. M., Pollak, S. D. (2019). Emotional Expressions Reconsidered: Challenges to Inferring Emotion From Human Facial Movements. https://journals.sagepub.com/doi/10.1177/1529100619832930

Benaissa, B., Koppen, M., Yoshida, K. (2018). Activity and Emotion Recognition for Elderly Health Monitoring. https://www.jstage.jst.go.jp/article/ijae/17/2/17IJAE-D-17-00020/article

Beyond Blue. (n.d.). Retrieved September 8, 2021, from https://www.beyondblue.org.au/media/statistics

Boustani, M., Schubert, C., Sennour, Y. (2007). The challenge of supporting care for dementia in primary care. Clinical Interventions in Aging, 2(4), 631‚Äì636.

CDC. (2021, January 6). Depression is Not a Normal Part of Growing Older | Alzheimer's Disease and Healthy Aging | CDC. https://www.cdc.gov/aging/depression/index.html

Chan, M., Esteve, D., Fourniols, J.-Y., Escriba, C., Campo, E. (2012). Smart wearable systems: Current status and future challenges. Artificial Intelligence in Medicine, 56(3), 137-156. https://doi.org/10.1016/j.artmed.2012.09.003

Chen, C., Johnson, J. G., Charles, K., Lee, A., Lifset, E. T., Hogarth, M., Moore, A. A., Farcas, E., Weibel, N. (2021). Understanding Barriers and Design Opportunities to Improve Healthcare and QOL for Older Adults through Voice Assistants. The

23rd International ACM SIGACCESS Conference on Computers and Accessibility, 1-16. https://doi.org/10.1145/3441852.3471218

Chen, Y., Shen, C. (2017). Performance Analysis of Smartphone-Sensor Behavior for Human Activity Recognition | IEEE Journals Magazine | IEEE Xplore. https://ieeexplore.ieee.org/document/7867876

Choi, Y.-J., Lee, Y.-W., Kim, B.-G. (2021). Residual-Based Graph Convolutional Network for Emotion Recognition in Conversation for Smart Internet of Things. Big Data, 9(4), 279-288. https://doi.org/10.1089/big.2020.0274

Christie, T. (n.d.). Human Identification in Video Surveillance Using Improved Gait Recognition Approaches [Qut research project]. Signal Processing, Artificial Intelligence and Vision Technologies. Retrieved May 9, 2022, from https://research.qut.edu.au/saivt/research-projects/human-identification-in-video-surveillance-using-improved-gait-recognition-approaches/

Chum, J., Kim, M. S., Zielinski, L., Bhatt, M., Chung, D., Yeung, S., Litke, K., McCabe, K., Whattam, J., Garrick, L., O'Neill, L., Goyert, S., Merrifield, C., Patel, Y., Samaan, Z. (2017). Acceptability of the Fitbit in behavioural activation therapy for depression: A qualitative study. Evidence-Based Mental Health, 20(4), 128-133. https://doi.org/10.1136/eb-2017-102763

Cook, J. D., Prairie, M. L., Plante, D. T. (2017). Utility of the Fitbit Flex to Evaluate Sleep in Major Depressive Disorder: A comparison against polysomnography and wrist-worn actigraphy. Journal of Affective Disorders, 217, 299-305. https://doi.org/

10.1016/j.jad.2017.04.030

Curry, D. (2020, October 23). Fitbit Revenue and Usage Statistics (2022). Business of Apps. https://www.businessofapps.com/data/fitbit-statistics/

De Nardin, L., Rodrigues, K. R. H., Zimmermann, L. C., Orlandi, B. D. M., da Grcaa C. Pimentel, M. (2020). Recognition of Human Activities via Wearable Sensors: Variables Identified in a Systematic Mapping. Proceedings of the Brazilian Symposium on Multimedia and the Web, 49-56. https://doi.org/10.1145/3428658.3431086

De Silva, L. C., Miyasato, T., Nakatsu, R. (1997). Facial emotion recognition using multi-modal information. Proceedings of ICICS, 1997 International Conference on Information, Communications and Signal Processing. Theme: Trends in Information Systems Engineering and Wireless Multimedia Communications (Cat., 1, 397-401 vol.1. https://doi.org/10.1109/ICICS.1997.647126

Depression and Older Adults. (n.d.). National Institute on Aging. Retrieved May 9, 2022, from https://www.nia.nih.gov/health/depression-and-older-adults

Domin, A., Spruijt-Metz, D., Theisen, D., Ouzzahra, Y., Vogele, C. (2021). Smartphone-Based Interventions for Physical Activity Promotion: Scoping Review of the Evidence Over the Last 10 Years. JMIR MHealth and UHealth, 9(7), e24308. https://doi.org/10.2196/24308

Drew, A. (2018). Bridging the Lab and the Real World. APS Observer, 31. https://www.psychologicalscience.org/observer/bridging-the-lab-and-the-real-world

Ekman, P. (2004). Emotional and Conversational Nonverbal Signals. In J. M. Larrazabal L. A. P. Miranda (Eds.), Language, Knowledge, and Representation (pp. 39,Äì50). Springer Netherlands. https://doi.org/10.1007/978-1-4020-2783-3$_3$

Elers, P., Hunter, I., Whiddett, D., Lockhart, C., Guesgen, H., Singh, A. (2018). User Requirements for Technology to Assist Aging in Place: Qualitative Study of Older People and Their Informal Support Networks. JMIR MHealth and UHealth, 6(6), e10741. https://doi.org/10.2196/10741

Garcia-Ceja, E., Riegler, M., Nordgreen, T., Jakobsen, P., Oedegaard, K. J., Torresen, J. (2018). Mental health monitoring with multimodal sensing and machine learning: A survey. Pervasive and Mobile Computing, 51, 1-26. https://doi.org/10.1016/j.pmcj.2018.09.003

Ghosh, A., Riccardi, G. (2014). Recognizing Human Activities from Smartphone Sensor Signals. Proceedings of the 22nd ACM International Conference on Multimedia, 865-868. https://doi.org/10.1145/2647868.2655034

Glisky, E. L. (2007). Changes in Cognitive Function in Human Aging. In D. R. Riddle (Ed.), Brain Aging: Models, Methods, and Mechanisms. CRC Press/Taylor Francis. http://www.ncbi.nlm.nih.gov/books/NBK3885/

Greenberg, S. A. (n.d.). The Geriatric Depression Scale (GDS).

Grunberg, J. (2010). "Pedogeriatrics": A pediatric nephrologist's outlook on common challenges facing pediatric and geriatric nephrologists. International Urology and

Nephrology, 42(1), 253‚Äì257. https://doi.org/10.1007/s11255-009-9593-2

Gu, F., Niu, J., He, Z., Jin, X., Rodrigues, J. J. P. C. (2017). SmartBuddy: An Integrated Mobile Sensing and Detecting System for Family Activities. GLOBECOM 2017 - 2017 IEEE Global Communications Conference, 1-7. https://doi.org/10.1109/GLOCOM.2017.8254140

Hamdy, R. C., Lewis, J. V., Kinser, A., Depelteau, A., Copeland, R., Kendall-Wilson, T., Whalen, K. (2017). Too Many Choices Confuse Patients With Dementia. https://journals.sagepub.com/doi/10.1177/2333721417720585

Han J, Richardson VE. The relationships among perceived discrimination, self-perceptions of aging, and depressive symptoms: a longitudinal examination of age discrimination. Aging Ment Health. 2015;19(8):747-55. doi: 10.1080/13607863.2014.962007. Epub 2014 Sep 30. PMID: 25266167.

Han, S., Williamson, B.D. Fong, Y. Improving random forest predictions in small datasets from two-phase sampling designs. BMC Med Inform Decis Mak 21, 322 (2021). https://doi.org/10.1186/s12911-021-01688-3.

Hartmann, R., Schmidt, F. M., Sander, C., Hegerl, U. (2019). Heart Rate Variability as Indicator of Clinical State in Depression. Frontiers in Psychiatry, 9, 735. https://doi.org/10.3389/fpsyt.2018.00735

Hasin, D., Link, B. (1988). Age and recognition of depression: Implications for a cohort effect in major depression. Psychological Medicine, 18(3), 683-688. https://doi.org/10.1017/S0033291700008369

Hernandez-Cruz, N., Garcia-Constantino, M., Beltran-Marquez, J., Cruz-Sandoval, D., Lopez-Nava, I. H., Cleland, I., Favela, J., Nugent, C., Ennis, A., Rafferty, J., Synnott, J. (2019). Study Design of an Environmental Smart Microphone System to Detect Anxiety in Patients with Dementia. Proceedings of the 13th EAI International Conference on Pervasive Computing Technologies for Healthcare, 383-388. https://doi.org/10.1145/3329189.3329234

Hsu, C.-C., Chien, Y. Y. (2009). An Intelligent Fuzzy Affective Computing System for Elderly Living Alone. 2009 Ninth International Conference on Hybrid Intelligent Systems, 1, 293-297. https://doi.org/10.1109/HIS.2009.318

Jamil, N., Mohd Khir, N. H., Ismail, M., Abdul Razak, F. H. (2015). Gait-Based Emotion Detection of Children with Autism Spectrum Disorders: A Preliminary Investigation.

Kadir, R. S. S. A., Murat, Z. H., Sulaiman, M. Z., Taib, M. N., Hanapiah, F. A., Omar, W. R. W. (2014). The preliminary investigation of electromagnetics radiation for the left hemisphere stroke. 2014 IEEE International Conference on Control System, Computing and Engineering (ICCSCE 2014), 606-610. https://doi.org/10.1109/ICCSCE.2014.7072790

Kasteren, Y. van, Bradford, D., Zhang, Q., Karunanithi, M., Ding, H. (2017). Understanding Smart Home Sensor Data for Ageing in Place Through Everyday Household Routines: A Mixed Method Case Study. JMIR MHealth and UHealth, 5(6), e5773. https://doi.org/10.2196/mhealth.5773

Kheirkhahan, M., Nair, S., Davoudi, A., Rashidi, P., Wanigatunga, A. A., Corbett,

D. B., Mendoza, T., Manini, T. M., Ranka, S. (2019). A smartwatch-based framework for real-time and online assessment and mobility monitoring. Journal of Biomedical Informatics, 89, 29-40. https://doi.org/10.1016/j.jbi.2018.11.003

Khundaqji, H., Hing, W., Furness, J., Climstein, M. (2020). Smart Shirts for Monitoring Physiological Parameters: Scoping Review. JMIR MHealth and UHealth, 8(5), e18092. https://doi.org/10.2196/18092

Ko, B. C. (2018). A Brief Review of Facial Emotion Recognition Based on Visual Information. Sensors, 18(2), 401. https://doi.org/10.3390/s18020401

Kone, C., Belleudy, C., Broutart, J. C., Le Thanh, N. (2017). Emotion recognition using musical mnemotherapy with three physiological sensors. In elsevier (Ed.), 10th International Conference on ICT, Society and Human Beings (p. 6). https://hal.archives-ouvertes.fr/hal-01624021

Koo, B. M., Vizer, L. M. (2019). Examining Mobile Technologies to Support Older Adults With Dementia Through the Lens of Personhood and Human Needs: Scoping Review. JMIR MHealth and UHealth, 7(11), e15122. https://doi.org/10.2196/15122

Kumar, P., Chauhan, R., Stephan, T., Shankar, A., Thakur, S. (2021). A Machine Learning Implementation for Mental Health Care. Application: Smart Watch for Depression Detection. 2021 11th International Conference on Cloud Computing, Data Science Engineering (Confluence), 568-574. https://doi.org/10.1109/Confluence51648.2021.9377199

Kyamakya, K., Al-Machot, F., Haj Mosa, A., Bouchachia, H., Chedjou, J. C., Bagula,

A. (2021). Emotion and Stress Recognition Related Sensors and Machine Learning Technologies. Sensors (Basel, Switzerland), 21(7), 2273. https://doi.org/10.3390/s21072273

Laricchia, F. (n.d.). Global smartwatch shipment forecast 2025. Statista. Retrieved April 26, 2022, from https://www.statista.com/statistics/878144/worldwide-smart-wristwear-shipments-forecast/

Laricchia, F. (2019). Global connected wearable devices 2016-2022. Statista. https://www.statista.com/statistics/487291/global-connected-wearable-devices/

Leber, J. (2014). An Atlas Of The Human Body That Maps Where We Feel Emotions. https://www.fastcompany.com/3024327/an-atlas-of-the-human-body-that-maps-where-we-feel-emotions

Li, K. H. C., White, F. A., Tipoe, T., Liu, T., Wong, M. C., Jesuthasan, A., Baranchuk, A., Tse, G., Yan, B. P. (2019). The Current State of Mobile Phone Apps for Monitoring Heart Rate, Heart Rate Variability, and Atrial Fibrillation: Narrative Review. JMIR MHealth and UHealth, 7(2), e11606. https://doi.org/10.2196/11606

Liu, Z., Heffernan, C., Tan, J. (2020). Caregiver burden: A concept analysis. International Journal of Nursing Sciences, 7(4), 438-445. https://doi.org/10.1016/j.ijnss.2020.07.012

Lustrek, M., Cvetkovifá, B., Mirchevska, V., Kafali, O., Romero, A. E., Stathis, K. (2015). Recognising lifestyle activities of diabetic patients with a smartphone. 2015 9th International Conference on Pervasive Computing Technologies for Healthcare (PervasiveHealth), 317-324. https://doi.org/10.4108/icst.pervasivehealth.2015.259118

Luxton, D. D., McCann, R. A., Bush, N. E., Mishkind, M. C., Reger, G. M. (n.d.). mHealth for mental health: Integrating smartphone technology in behavioral healthcare. - PsycNET. Retrieved February 17, 2022, from https://psycnet.apa.org/doiLanding?doi=10.1037%2Fa0024485

Lyons, E. J., Swartz, M. C., Lewis, Z. H., Martinez, E., Jennings, K. (2017). Feasibility and Acceptability of a Wearable Technology Physical Activity Intervention With Telephone Counseling for Mid-Aged and Older Adults: A Randomized Controlled Pilot Trial. JMIR MHealth and UHealth, 5(3), e6967. https://doi.org/10.2196/mhealth.6967

Miranda, J. A., Canabal, M. F., Gutierrez-Martin, L., Lanza-Gutierrez, J. M., Portela-Garcia, M., Lopez-Ongil, C. (2021). Fear Recognition for Women Using a Reduced Set of Physiological Signals. Sensors (Basel, Switzerland), 21(5), 1587. https://doi.org/10.3390/s21051587

Mughal, F., Raffe, W., Garcia, J. (2020). Emotion Recognition Techniques for Geriatric Users: A Snapshot. 2020 IEEE 8th International Conference on Serious Games and Applications for Health (SeGAH), 1,Äì8. https://doi.org/10.1109/SeGAH49190.2020.9201749

Mughal, F., Raffe, W., Stubbs, P., Garcia, J. (2021). Protocol: Autonomous Monitoring for Older Aged Persons. https://osf.io/gcyrm/

Mughal, F., Raffe, W., Stubbs, P., Kneebone, I., Garcia, J. (2021). Fitbits for monitoring depressive tendencies in older aged persons: Qualitative outcomes of a feasibility study (Preprint). https://opus.lib.uts.edu.au/handle/10453/154022

Murman, D. L. (2015). The Impact of Age on Cognition. Seminars in Hearing, 36(3), 111‚Äì121. https://doi.org/10.1055/s-0035-1555115

Naslund, J. A., Aschbrenner, K. A., Bartels, S. J. (2016). Wearable Devices and Smartphones for Activity Tracking Among People with Serious Mental Illness. Mental Health and Physical Activity, 10, 10‚Äì17. https://doi.org/10.1016/j.mhpa.2016.02.001

Naslund, J., Aschbrenner, K., Barre, L., Bartels, S. (2014). Feasibility of Popular m-Health Technologies for Activity Tracking Among Individuals with Serious Mental Illness. Telemedicine Journal and E-Health: The Official Journal of the American Telemedicine Association, 21. https://doi.org/10.1089/tmj.2014.0105

Nature. (2019). From the lab to real-world use. Nature Sustainability, 2(11), 989‚Äì989. https://doi.org/10.1038/s41893-019-0435-7 NIA. (2017). Supporting Older Patients with Chronic Conditions. National Institute on Aging. http://www.nia.nih.gov/health/supporting-older-patients-chronic-conditions

No Isolation. (2018). Why do many seniors have trouble using technology? No Isolation. https://www.noisolation.com/research/why-do-many-seniors-have-trouble-using-technology

Nummenmaa, L., Glerean, E., Hari, R., Hietanen, J. K. (2014). Bodily maps of emotions. Proceedings of the National Academy of Sciences, 111(2), 646-651. https://doi.org/10.1073/pnas.1321664111

Park, S., Constantinides, M., Aiello, L. M., Quercia, D., Van Gent, P. (2020). Well-Beat: A Framework for Tracking Daily Well-Being Using Smartwatches. IEEE Internet Computing, 24(5), 10-17. https://doi.org/10.1109/MIC.2020.3017867

Park, S., Kim, T., Shin, S. A., Kim, Y. K., Sohn, B. K., Park, H.-J., Youn, J.-H., Lee, J.-Y. (2017). Behavioral and Neuroimaging Evidence for Facial Emotion Recognition in Elderly Korean Adults with Mild Cognitive Impairment, Alzheimer's Disease, and Frontotemporal Dementia. Frontiers in Aging Neuroscience, 9, 389. https://doi.org/10.3389/fnagi.2017.00389

Patlar Akbulut, F., Ikitimur, B., Akan, A. (2020). Wearable sensor-based evaluation of psychosocial stress in patients with metabolic syndrome. Artificial Intelligence in Medicine, 104, 101824. https://doi.org/10.1016/j.artmed.2020.101824

Pedrelli P, Fedor S, Ghandeharioun A, Howe E, Ionescu DF, Bhathena D, Fisher LB, Cusin C, Nyer M, Yeung A, Sangermano L, Mischoulon D, Alpert JE, Picard RW. Monitoring Changes in Depression Severity Using Wearable and Mobile Sensors. Front Psychiatry. 2020 Dec 18;11:584711. doi: 10.3389/fpsyt.2020.584711. PMID: 33391050; PMCID: PMC7775362.

Pollreisz, D., TaheriNejad, N. (2017). A simple algorithm for emotion recognition, using physiological signals of a smart watch. 2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), 2353,Äì2356. https://doi.org/10.1109/EMBC.2017.8037328

Pratap, A., Atkins, D. C., Renn, B. N., Tanana, M. J., Mooney, S. D., Anguera, J. A., Arean, P. A. (2019). The accuracy of passive phone sensors in predicting daily mood. Depression and Anxiety, 36(1), 72-81. https://doi.org/10.1002/da.22822

Quiroz, J. C., Geangu, E., Yong, M. H. (2018). Emotion Recognition Using Smart Watch Sensor Data: Mixed-Design Study. JMIR Mental Health, 5(3), e10153. https://doi.org/10.2196/10153

Randhavane, T., Bhattacharya, U., Kapsaskis, K., Gray, K., Bera, A., Manocha, D. (2020). Identifying Emotions from Walking using Affective and Deep Features. ArXiv:1906.11884 [Cs]. http://arxiv.org/abs/1906.11884

Ridgers, N. D., McNarry, M. A., Mackintosh, K. A. (2016). Feasibility and Effectiveness of Using Wearable Activity Trackers in Youth: A Systematic Review. JMIR MHealth and UHealth, 4(4), e129. https://doi.org/10.2196/mhealth.6540

Ringeval, M., Wagner, G., Denford, J., Pare, G., Kitsiou, S. (2020). Fitbit-Based Interventions for Healthy Lifestyle Outcomes: Systematic Review and Meta-Analysis. Journal of Medical Internet Research, 22(10), e23954. https://doi.org/10.2196/23954

Rohani, D. A., Faurholt-Jepsen, M., Kessing, L. V., Bardram, J. E. (2018). Correlations Between Objective Behavioral Features Collected From Mobile and Wearable Devices and Depressive Mood Symptoms in Patients With Affective Disorders: Systematic Review. JMIR MHealth and UHealth, 6(8), e9691. https://doi.org/10.2196/mhealth.9691

Rozanska, A., Rachwaniec-Szczecinska, Z., Kawala-Janik, A., Podpora, M. (2018). Internet of Things Embedded System for Emotion Recognition. 2018 IEEE 20th International Conference on E-Health Networking, Applications and Services (Healthcom), 1,Äì5. https://doi.org/10.1109/HealthCom.2018.8531100

Rykov Y., Thach TQ, Bojic I, Christopoulos G, Car J. Digital Biomarkers for Depression Screening With Wearable Devices: Cross-sectional Study With Machine Learning Modeling. JMIR Mhealth Uhealth. 2021 Oct 25;9(10):e24872. doi: 10.2196/24872. PMID: 34694233; PMCID: PMC8576601.

Saeb, S., Zhang, M., Kwasny, M., Karr, C., Kording, K., Mohr, D. (2015, August 3). The Relationship between Clinical, Momentary, and Sensor-based Assessment of Depression. 9th International Conference on Pervasive Computing Technologies for Healthcare. https://eudl.eu/doi/10.4108/icst.pervasivehealth.2015.259034

Sahu, K. S., Oetomo, A., Morita, P. P. (2020). Enabling Remote Patient Monitoring Through the Use of Smart Thermostat Data in Canada: Exploratory Study. JMIR MHealth and UHealth, 8(11), e21016. https://doi.org/10.2196/21016

Sas, C., Brahney, K., Oechsner, C., Trivedi, A., Nomesque, M., Mughal, Z., Cheverst, K. W., Clinch, S., Davies, N. (2017). Communication Needs of Elderly at Risk of Falls and their Remote Family. Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems, 2900-2908. https://doi.org/10.1145/3027063.3053274

Sathyanarayana, A., Joty, S., Fernandez-Luque, L., Ofli, F., Srivastava, J., Elmagarmid, A., Arora, T., Taheri, S. (2016). Sleep Quality Prediction From Wearable Data Using Deep Learning. JMIR MHealth and UHealth, 4(4), e6562. https://doi.org/10.2196/mhealth.6562

Schneider, S., Christensen, A., Haubinger, F. B., Fallgatter, A. J., Giese, M. A., Ehlis, A.-C. (2014). Show me how you walk and I tell you how you feel‚ÄîA functional near-infrared spectroscopy study on emotion perception based on human gait. NeuroImage, 85 Pt 1, 380‚Äì390. https://doi.org/10.1016/j.neuroimage.2013.07.078

Schober, P., Boer, C., Schwarte, L. (2018). Correlation Coefficients: Appropriate Use and Interpretation: Anesthesia Analgesia.

scikit-learn. (n.d.-a). 3.1. Cross-validation: Evaluating estimator performance. Scikit-Learn. Retrieved March 18, 2022, from `https://scikit-learn/stable/modules/cross_validation.html`

scikit-learn. (n.d.-b). scikit-learn: Machine learning in Python. Retrieved May 10, 2022, from https://scikit-learn.org/stable/

Segal, D. L., Coolidge, F. L., Mincic, M. S., O'Riley, A. (2005). Beliefs about mental illness and willingness to seek help: A cross-sectional study. Aging Mental Health, 9(4), 363-367. https://doi.org/10.1080/13607860500131047

Servia-Rodriguez, S., Rachuri, K. K., Mascolo, C., Rentfrow, P. J., Lathia, N., Sandstrom, G. M. (2017). Mobile Sensing at the Service of Mental Well-being: A Large-scale Longitudinal Study. Proceedings of the 26th International Conference on World Wide Web, 103‚Äì112. https://doi.org/10.1145/3038912.3052618

Sheikh, J. I., Yesavage, J. A. (1986). Geriatric Depression Scale (GDS): Recent evidence and development of a shorter version. Clinical Gerontologist: The Journal of Aging and Mental Health, 5(1-2), 165‚Äì173. https://doi.org/10.1300/J018v05n01$_0$9

Shu, L., Xie, J., Yang, M., Li, Z., Li, Z., Liao, D., Xu, X., Yang, X. (2018). A Review of Emotion Recognition Using Physiological Signals‚ÄîPMC. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6069143/

Shu, L., Yu, Y., Chen, W., Hua, H., Li, Q., Jin, J., Xu, X. (2020). Wearable Emotion Recognition Using Heart Rate Data from a Smart Bracelet. Sensors, 20(3), 718. https://doi.org/10.3390/s20030718

Sivapalan, S. (2014). Human identification from video using advanced gait recognition techniques [Phd, Queensland University of Technology]. https://eprints.qut.edu.au/77620/

Soon, S., Svavarsdottir, H., Downey, C., Jayne, D. G. (2020). Wearable devices for remote vital signs monitoring in the outpatient setting: An overview of the field. BMJ Innovations, 6(2). https://doi.org/10.1136/bmjinnov-2019-000354

Sultana, M., Al-Jaefri, M., Lee, J. (2020). Using Machine Learning and Smartphone and Smartwatch Data to Detect Emotional States and Transitions: Exploratory Study. https://mhealth.jmir.org/2020/9/e17818/

Tong, H. L., Coiera, E., Laranjo, L. (2018). Using a Mobile Social Networking App to Promote Physical Activity: A Qualitative Study of Users' Perspectives. Journal of Medical Internet Research, 20(12), e11439. https://doi.org/10.2196/11439

Torgyn Shaikhina, Dave Lowe, Sunil Daga, David Briggs, Robert Higgins, Natasha Khovanova, Machine Learning for Predictive Modelling based on Small Data in Biomedical Engineering, IFAC-PapersOnLine, Volume 48, Issue 20, 2015, Pages 469-474, ISSN 2405-8963, https://doi.org/10.1016/j.ifacol.2015.10.185.

Vaportzis, E., Giatsi Clausen, M., Gow, A. J. (2017). Older Adults Perceptions of Technology and Barriers to Interacting with Tablet Computers: A Focus Group Study. Frontiers in Psychology, 8, 1687. https://doi.org/10.3389/fpsyg.2017.01687

Virtanen, M., Singh-Manoux, A., Batty, G. D., Ebmeier, K. P., Jokela, M., Harmer, C. J., Kivimaki, M. (2017). The level of cognitive function and recognition of emotions in older adults. PLoS ONE, 12(10), e0185513. https://doi.org/10.1371/journal.pone.0185513

Voit, A., Stowell, E., Weber, D., Witte, C., Karcher, D., Henze, N. (2016). Envisioning an ambient smart calendar to support aging in place. Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct, 1596‚Äì1601. https://doi.org/10.1145/2968219.2968555

Wang, S., He, M., Gao, Z., He, S., Ji, Q. (2014). Emotion recognition from thermal infrared images using deep Boltzmann machine. Frontiers of Computer Science, 8(4), 609-618. https://doi.org/10.1007/s11704-014-3295-3

WHO. (2017). Mental health of older adults. https://www.who.int/news-room/fact-sheets/detail/mental-health-of-older-adults

WHO. (2021). Depression. https://www.who.int/news-room/fact-sheets/detail/depression

Wiechetek Ostos, M., Schenk, F., Baenziger, T., von Gunten, A. (2011). An exploratory study on facial emotion recognition capacity in beginning Alzheimer's disease. European Neurology, 65(6), 361-367. https://doi.org/10.1159/000327979

Xiaoxi, M., Weisi, L., Dongyan, H., Minghui, D., Li, H. (2017). Facial emotion recognition. 2017 IEEE 2nd International Conference on Signal and Image Processing (ICSIP), 77-81. https://doi.org/10.1109/SIPROCESS.2017.8124509

Xu, W., Zammit, K. (2020). Applying Thematic Analysis to Education: A Hybrid Approach to Interpreting Data in Practitioner Research,ÄîWen Xu, Katina Zammit, 2020. https://journals.sagepub.com/doi/full/10.1177/1609406920918810

Yaddaden, Y., Bouzouane, A., Adda, M., Bouchard, B. (2016). A New Approach of Facial Expression Recognition for Ambient Assisted Living | Proceedings of the 9th ACM International Conference on PErvasive Technologies Related to Assistive Environments. https://dl.acm.org/doi/10.1145/2910674.2910703

Zangeneh Soroush, M., Maghooli, K., Setarehdan, S. K., Motie Nasrabadi, A. (2017). A Review on EEG Signals Based Emotion Recognition. International Clinical Neuroscience Journal, 4(4), 118,Äì129. https://doi.org/10.15171/icnj.2017.01

Zhang, Z., Song, Y., Cui, L., Liu, X., Zhu, T. (2016). Emotion recognition based on customized smart bracelet with built-in accelerometer. PeerJ, 4, e2258. https://doi.org/10.7717/peerj.2258

Zhao, M., Adib, F., Katabi, D. (2018). Emotion recognition using wireless signals | Communications of the ACM. https://dl.acm.org/doi/10.1145/3236621

Zimmermann, L. C., Rodrigues, K. R. da H., Pimentel, M. da G. C. (2019). EPARS: Elderly physical activity reminder system using smartphone and wearable sensors. Adjunct Proceedings of the 2019 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2019 ACM International Symposium on Wearable Computers, 1139-1145. https://doi.org/10.1145/3341162.3350845