A MOBILE SENSING APP FOR MENTAL HEALTH TO SUPPORT FEDERATED LEARNING

by

Banuchitra Suruliraj

Submitted in partial fulfilment of the requirements
for the degree of Master of Computer Science

at

Dalhousie University
Halifax, Nova Scotia
December 2020

DEDICATION

This thesis is dedicated to my father

Late Mr. Suruliraj Ramalingam

He taught me to dream beyond the glass ceiling

"Daddy, one day, I will make you proud"

TABLE OF CONTENTS

DEDICATION ... **ii**

TABLE OF CONTENTS ... **iii**

LIST OF TABLES ... **viii**

LIST OF FIGURES ... **x**

ABSTRACT ... **xii**

LIST OF ABBREVIATIONS USED ... **xiii**

ACKNOWLEDGEMENTS ... **xvi**

CHAPTER 1 INTRODUCTION ... **1**

1.1 The Problem ... **1**

1.2 Motivation ... **3**

1.3 Solution ... **4**

1.4 Contributions ... **5**

1.5 Overview of Thesis ... **5**

CHAPTER 2 RESEARCH BACKGROUND ... **7**

2.1 Mental Health Monitoring mobile applications ... **7**

2.1.1 Paper Inclusion Process ... 7

2.1.2 Targeted mental health aspects ... 9

2.1.3 Mental health illness ... 10

2.1.4 Mental health Indicators ... 12

2.1.5 Sensing types in mobile apps ... 14

2.1.6 Mobile sensing apps for Detection ... 16

2.1.7 Mobile sensing apps for Association 19
2.1.8 Mobile sensing apps for Intervention 23
2.2 Discussion of Related Work 24
2.2.1 Mobile sensing applications by operating systems 24
2.2.2 Mobile sensing apps by country or region 25
2.2.3 Mobile sensing applications for mental health evaluation details 26
2.3 Summary of gaps in the existing literature 30
2.4 Mobile sensing application for Mental health with Federated Learning 31
2.4.1 Federated learning overview 31
CHAPTER 3 PROSIT DESIGN AND IMPLEMENTATION 36
3.1 Early Development Process 36
3.2 Domain Expertise – Psychiatry 38
3.3 PROSIT DESIGN 39
3.3.1 The Architecture 40
3.3.2 Design decisions for application configurations 43
3.3.3 Client-side process flow 46
3.3.4 Privacy and Security 48
3.3.5 Power optimization 49
3.3.6 Memory optimization 49
3.3.7 Security 50
3.3.8 Federated Learning for Mental Health Monitoring Systems 52
CHAPTER 4 PROSIT EVALUATION 58

4.1 Research questions ... 58

4.1.1 Approaches ... 59

4.2 Online Survey ... 60

4.2.1 Survey/Questionnaire Design ... 62

4.2.2 Answering the R1: ... 62

4.2.3 Answering R2 ... 64

4.2.4 Answering R3 and R4 ... 67

4.2.5 Answering R5 ... 67

4.3 Online Survey ... 68

4.3.1 Tools used ... 68

4.3.2 Participants demographics Mturk Study ... 68

4.4 Pilot Study for App Functionality Evaluation ... 75

4.5 Anomaly detection feasibility Study ... 75

4.5.1 Client-side anomaly detection algorithm implementation ... 75

4.5.2 Performance evaluation of PROSIT and PROSITLite ... 76

4.5.3 Performance evaluation data collection ... 76

CHAPTER 5 STUDY RESULTS ... 77

5.1 Online survey ... 77

5.1.1 Tracking Fine-grained Versus Summary Tracking ... 77

5.1.2 Feature ratings ... 78

5.1.3 Factors affecting user perception ... 81

5.2 PROSIT app pilot study ... 84

5.2.1 Dataset information 84

5.2.2 Pre-processing 85

5.2.3 Exploratory Data Analysis 86

5.2.4 Principal Component Analysis 88

5.2.5 Anomaly detection algorithm 91

5.2.6 Power consumption and Memory usage 91

5.3 Discussions of the Results 92

5.3.1 Privacy score of Mobile Sensing features 93

5.3.2 Factors affecting privacy perception 94

5.3.3 PROSIT as Mental Health Monitoring System 97

5.3.4 Federated Learning framework for detecting Depression 99

5.3.5 Clinical Blurb from Psychiatrist 100

5.3.6 Future work in Anomaly detection and Federated Learning 101

CHAPTER 6 CONCLUSION 102

6.1 Limitations 102

6.2 Contributions 103

6.3 Future work 103

6.4 Conclusion 103

BIBLIOGRAPHY 105

APPENDIX A. Privacy Perceptions scales 124

APPENDIX B. Permission to Use 127

APPENDIX C. Thesis Approval Form 128

APPENDIX D. Medium fidelity prototype PROSIT app 129

APPENDIX E. Screenshots of PROSIT app... 131

APPENDIX F. Mturk Survey Questions 135

APPENDIX G. PROSIT Team consent.. 151

APPENDIX H. Research Ethics Board Approval Letter................................... 152

APPENDIX I. PROSIT Dataset usage approval.. 155

APPENDIX J. Agile process timeline 156

APPENDIX K. Publications During My Master's Degree 160

APPENDIX L. Publications Related to My Thesis 161

LIST OF TABLES

Table 3.1 Sensing types included in PROSIT iOS and Android Versions 36

Table 3.2 Local table structure... 52

Table 3.3 Global table structure 53

Table 4.1 Research Questions and their targeted outcomes 59

Table 4.2 Sample question on the comparison between Raw data vs Summary............. 63

Table 4.3 Core scales for factors affecting privacy perception 64

Table 4.4 Participant demographics... 74

Table 5.1 Descriptive Statistics for the two modes of tracking 77

Table 5.2 Paired sample T-test for Fine-grained and Summary tracking 77

Table 5.3 Feature rating questions for tracking All the time versus Summary 78

Table 5.4 Descriptive statistics for feature ratings for tracking Fine-grained data tracking versus Summary data tracking 79

Table 5.5 Paired-Sample test feature ratings for tracking Summary data tracking versus Fine-grained data tracking 80

Table 5.6 Factors affecting privacy perceptions 81

Table 5.7 Correlation between privacy factors and feature preferences.......................... 82

Table 5.8 Feature description... 85

Table 5.9 Feature ranking and importance scores 90

Table 5.10 Comparison of PROSIT and PROSITLite... 92

Table 5.11 Mean rating and Standard deviation for features.. 93

Table 5.12 Factors affecting mobile sensing feature ratings - Correlation...................... 95

Table 5.13 Correlation analysis between features and ground truth................................ 98

LIST OF FIGURES

Figure 2.1 - Paper inclusion process 9

Figure.2.2 Occurrence of targeted mental health conditions and indicators 10

Figure.2.3 Sensing types and smartphone feasibility 15

Figure 2.4 Detection - Mental condition and number of studies 16

Figure 2.5 Association - Mental condition and number of studies 20

Figure 2.6 Mobile sensing applications for mental health by the operating system 25

Figure 2.7 Mobile sensing application research by Country or region 26

Figure 2.8 Mobile sensing apps by participant type 28

Figure 2.9 Mobile sensing apps by Study Method 28

Figure 2.10 Percentage of research that took the listed measures into account 29

Figure 2.11 Next word predictions in Gboard. Based on the con-text "I love you", the keyboard predicts "and", "too", and "so much" [59] 32

Figure 2.12 An illustration of the federated learning process from [89]: (A) client devices compute Stochastic Gradient Descent (SGD) values and updates on locally-stored data, (B) a server aggregates the client updates to build a new global model, (C) the new model is sent back to clients, and the process is repeated 33

Figure 2.13 Federated Learning Protocol [14] 34

Figure 3.1 Agile development weekly sprint 39

Figure 3.2 PROSIT iOS version - a) Login screen b) Home screen 40

Figure 3.3 PROSIT System Architecture 41

Figure 3.4 High-level overview of app process flow 47

Figure 3.5 Login sequence diagram 50

Figure 3.6 Upload sequence diagram 51

Figure 3.7 PROSIT architecture - With on-device learning 57

Figure 4.1 - Processes of inclusion in Online survey 61

Figure 4.2 Depiction of sensing types in the PROSIT app 63

Figure 4.3 Demographics by gender 69

Figure 4.4 Demographics by country 70

Figure 4.5 Demographics by age 71

Figure 4.6 Demographics by education 71

Figure 4.7 Demographics by the history of mental illness 72

Figure 4.8 Demographics by past mobile app usage for mental health 72

Figure 4.9 Demographics by fitness app use 73

Figure 4.10 Demographics by mobile device 74

Figure 5.1 Correlation heatmap 87

Figure 5.2 PCA Explained variance ratio 88

Figure 5.3 K-nearest neighbor classifier result 3D scatter plot 89

Figure 5.4 Extra Trees Classifier feature importance 90

Figure 5.5 Average ratings of tracking features 94

Figure 5.6 Factors affecting mobile sensing feature privacy comfortability ratings 96

Figure 5.7 Battery consumption of PROSIT and PROSITLite 99

Figure 5.8 Memory consumption PROSIT and PROSITLite 99

ABSTRACT

Smartphones are used by half of the world population. More than 10,000 applications are targeted at Mental health. Available apps are limited in four major ways: One, most apps are designed for the Android platform, 80% of the apps did not consider studying both iOS and Android users. Two, there is a lack of a comprehensive tool to study multiple mental health issues. Three, although these apps collect privacy-sensitive data, 67% of studies did not take the privacy concerns of the users into account. Finally, there is an overhead in terms of battery, internet, storage, and time in centralized data analysis.

To overcome the limitations, in this thesis we present the design, development, and field evaluation of two mobile sensing applications called *PROSIT* and *PROSITLite. PROSIT* app passively and unobtrusively collects mobile sensor data and periodically transfers the data to secure servers. The app tracks 23 different sensor data and hence serves as a comprehensive tool to study different mental health issues. The app runs on both iOS and Android platforms, thus, accessible to over 98% of smartphone users. We conducted an online survey to evaluate the users' comfortability with *PROSIT* and privacy concerns, the results from 491 participants show that users are comfortable to track all the app features. Perceptions about surveillance, intrusion, and data leakage influence users' comfortability negatively whereas trust, control, and consent have a positive influence in user comfortability. We conducted a pilot study on 18 participants who used the app for 2 weeks. The results of the Principal Component Analysis and K-Nearest Neighbours classifier show 73% accuracy in distinguishing the patients and non-patients. Further, we propose a Federated Learning (FL) framework for mental health monitoring to overcome the overheads and preserve privacy. To lay a foundation for FL, we developed *PROSITLite* with an anomaly detection algorithm to detect Depression. Results from the feasibility study show *PROSITLite* is efficient in overcoming the identified overheads. In the future, we aim to train a robust model, with data from the ongoing study and implement on-device training with full implementation of federated learning.

LIST OF ABBREVIATIONS USED

Amazon Mturk	Amazon Mechanical Turk
ANOVA	Analysis of Variance
APC	App Permission Concerns
API	Applications Programming Interface
BI	Behavioral Intention
CA	Computer Anxiety
CoreML	Core Machine Learning
COVID-19	Corona Virus Disease - 2019
dB	Decibel
DDoS	Distributed Denial of Service
EMA	Electronic Momentary Assessment
Gboard	Google virtual keyboard
GPS	Global Positioning System
GSR	Galvanic Skin Resistance
HCI	Human-Computer Interaction
HIT	Human Intelligence Tasks
HRV	Heart Rate Variability
HTTP	HyperText Transfer Protocol
INT	Intention to accept app permissions
iOS	iPhone Operating System

JSON	JavaScript Object Notation
KB	Kilo Bytes
KNN	K-Nearest Neighbours
MAD	Mean Absolute Deviation
MEC	Mobile Edge Computing
MHMS	Mental Health Monitoring Systems
MUIPC	Mobile Users Information Privacy Concerns Instrument
NASA	National Aeronautics and Space Administration
OCR	Optical Character Recognition
OAEP	Optimal asymmetric encryption padding
PB	Perceived Benefit
PC	Perceived Control
PCA	Principal Component Analysis
PI	Perceived Intrusion
PPB	Privacy Protection Behavior
PPE	Past Privacy Experience
PPG	Photoplethysmography
PRC	Privacy Concern
PROSIT	Predicting Risks and Outcomes of Social InTeractions

PS	Perceived Surveillance
R1	Research Question 1
R2	Research Question 2
R3	Research Question 3
R4	Research Question 4
R5	Research Question 5
REST	Representational State Transfer
RSA	Rivest–Shamir–Adleman
SGD	Stochastic Gradient Descent
SIM	Subscriber Identification Module
SMI	Serious Mental Illnesses
SU	Secondary Use of personal information
TR	Trust in app developers and health care providers
UNDESA	United Nations Department of Economic and Social Affairs
Wi-Fi	Wireless Fidelity

ACKNOWLEDGEMENTS

Firstly, I say a big "Thank You" to my master's degree supervisor, Dr. Rita Orji for her tremendous support, guidance, and suggestions during this work. Thank you, Dr. Meier Sandra, for accepting me as one of your research assistants for these two years. I would have not been able to carry out this research without the funding I was provided. The funds kept me away from doing odd jobs in night shifts. To both my supervisors, thank you for believing in me and being patient with my mistakes.

I also thank my committee members, Dr. Sandra Meier and Dr. Derek Reilly. Thank you for volunteering to patiently read my thesis and provide valuable feedback.

Thanks to Dal Ph.D. students Felwah Alqahtani, Chinenye Ndulue, and Oladapo Oyebode for clarifying my trivial questions and helping me throughout my program.

I especially want to thank my husband who sacrificed part of his career to take care of our little son, while I focused on my education. I'm deeply indebted to my exceptional parent-in-laws Mr. Rajasekaran and Mrs. Leelavathi who spent their entire savings to sponsor my higher education, thank you for all the love and care.

I thank my father in heaven Mr. Suruliraj, who could not live to see me graduate. He defended me against the Indian social norms for women. My sincere thanks to my mother Mrs. Indira for being a role model of a strong woman, what I'm today is all because of you both.

Above all, I thank my 3-year-old son, for adapting to all the harsh changes imposed by me pursuing a master's degree.

Finally, I thank the Almighty for giving me strength and making this dream a reality!

CHAPTER 1 INTRODUCTION

1.1 The Problem

The recent years have witnessed tremendous advances in technology especially in the area of mobile and handheld devices which has become part of our day-to-day lives. According to recent statistics, 3.5 billion people worldwide own a smartphone [3]. It is among the popular devices with high penetration and adoption across all demography, age groups, and culture inclusive [121]. 81% of Americans and 86% of Canadians own a smartphone.

Considering this increasing adoption, coupled with increased processing power, recent years have witnessed increasing investment into designing and developing mobile applications to solve problems in many domains including E-commerce, Hospitality, Food, Entertainment, and Healthcare [102]. In the area of Healthcare, several applications have been developed to support various health and wellness objectives such as doctor-on-demand, medical tracking, healthy lifestyle, dieting, women's health, and mental health. As of 2017, there were approximately 325,000 health and wellness apps [4], with more than 10,000 apps designed specifically for mental health [135].

Mental Health issues are one of the leading causes of disability [161]. The economic burden of these disorders is huge, in terms of economic costs, reduced workforce participation, occupational impairment, and lost productivity [32][54][13], they also have an indirect effect on life expectancy [72], People with a serious mental illness (SMI) die 25 years earlier than those unaffected [140]. Almost a third of countries (31%) do not have an allocated budget for mental health at all compared to those countries that do [118]. According to the World Health Organization (WHO) report, approximately 320 million people are suffering from depression, which is almost 4.4% of the world's entire population [171].

Many mental health disorders are chronic and relapsing in nature, and hence require long-term follow up and assessment. Traditional monitoring methods rely on retrospective reports which are subject to recall bias. Mental health is an area of healthcare that still largely relies on traditional and subjective methods of diagnosis performed using checklist type criteria as stated by Tuso [136], This approach limits the ability to accurately understand behavior change in real-world settings [120].

Currently, there are 3.5 billion smartphone users in the world [3]. People keep their phones with or near them and use them frequently. On average, people check their phones 46 times per day, and for younger people, that figure is 85 times per day [6]. Smartphones have a variety of embedded sensors. These can include communication devices (Wi-Fi, Bluetooth, etc.), inertial sensors (accelerometer, gyroscope, etc.), and can be effectively used to track location changes and communication patterns. Additionally, if a person owns a wearable device like a smartwatch or fitness bands, all the sensing information from the wearables can be collected and stored in a synchronized smartphone. This opens the possibility of mobile sensing applications in the healthcare domain [34]. Mobile phones help to provide better treatment to more users with less cost, particularly in the area of mental health care [53]. The earliest study in 2010 [111] involved using a sensing device to be carried by the participants all the time. An important sub-field that has been rising in the past years is the application of mHealth technology in the field of mental and behavioral health, enabling researchers to study stress, depression, mood, personality change, schizophrenia, physical activity, and addictive behavior, among other things [95]. An increasing number of studies demonstrate the ability of behavior-monitoring devices to assist in addressing the challenges of self-tracking. The technical barriers are related to performing privacy-sensitive and resource-sensitive reasoning with noisy data and noisy labels and providing useful and effective feedback to users [71]. This field is growing over the past decade, with the majority of the works focused on Depression, Bipolar, Schizophrenia, Mood Disorders, and Post Traumatic Stress Disorder (PTSD) [22]. A significant amount of research links behavior to mental health [37]. Current challenges include issues such as battery consumption vs data collection strategies, real-time vs. batch data analysis, data storage challenges, and data security and privacy [51].

Collaboration plays a key role in interdisciplinary research. Behavioral sensing technologies, if developed closely with patients and their clinicians, will be accepted more widely in clinical practice and provide more effective management of Serious Mental Illnesses (SMI) [85]. Design, development, and deployment of mobile sensing technologies for mental health will entail close collaboration between professionals in Human-Computer Interaction (HCI) and mental health care backgrounds and user groups [38].

Personal sensing is still in its infancy, it holds great promise for monitoring at-risk populations and providing the foundation for the next generation of mobile health (or mHealth) interventions [96]. Mobile sensing applications for mental health passively and continuously monitor and transfer bulk data to computational systems serving as backend. These applications often rely on background server or cloud-based services for emotion inference. As a result, these approaches suffer from privacy concerns and network latency [144]. Smartphones are increasingly equipped with more memory and computational power. In 2012, NASA stated that a consumer smartphone had 100 times more computing power than an average satellite [101]. Despite the increasing power, processing mobile sensing data in smartphones is still underexplored by researchers.

1.2 Motivation

Stigma towards and discrimination against people with mental health issues is a significant barrier to mental health service utilization [127]. This results in several million people with mental issues going about untreated. Developing countries spend less than 1% of their small health budgets on mental health services [118], roughly 40% of people suffering from depression live in these developing nations [171]. With the advent of smartphones, finding their widespread use across all populations, digital services for mental health delivered through smartphones will significantly reduce the health care cost and provide better treatment options which in turn will be beneficial to people in the developing nations especially in countries where mental health is stigmatized.

Studies have proven that mental health monitoring systems can be established from the day-to-day information of the physiological states of individuals says INSIGHT [83]. A scoping review published in 2020 states that research is needed on ways to customize implementation and evaluation of mobile applications for mental health to ensure skill development to provide quality of care [62]. Another review of 35 literature articles states that despite significant research in this area, the applications are not yet used in the health care system because most studies described their methods as trials, gathered data from a small sample size [65]. Hence, we wanted to identify significant gaps in this research domain and bridge them with best practices from two disciplines Computer science and Psychiatry.

There are four major gaps in the area of MHMS, firstly most of the mobile sensing research is conducted using smartphones running on the Android operating system, whereas 54% of North American users use iOS. This implies that more than half of the population is marginalized. Secondly, mobile sensing applications raise privacy concerns among users since the app collects and transfers bulk data to the backend [144]. Thirdly the method of transferring data to centralized servers and analysis result pose significant time, storage, power, and network overheads on both client and server-side. Finally, there is a lack of one comprehensive tool for clinicians to study multiple mental health issues. Most apps track an average of 4 different features and study a maximum of three mental health issues.

1.3 Solution

To bridge the gaps in this domain, this research presents the design, development, field evaluation of a mobile sensing application called "PROSIT". The app designed to run on iOS and Android platforms continuously collects more than 23 different sensor data about the user's daily life through smartphone sensors and periodically transfers the information to secure servers. This information will be used by the psychiatrists to predict mental health incidences and gain day-to-day insights about their patients' mental health, associated factors, and possible treatment responses. To investigate the privacy concerns related to using mobile sensing app for mental health-related purposes, we conducted an online survey of 491 participants to understand the user perceptions about various mobile sensing features and associated factors. The results from the survey show that user perceptions about surveillance, intrusion, and data leakage are factors that make the users uncomfortable about the tracking features of the app. Trust, control, and consent are the factors that make users comfortable with using mobile sensing app features.

Next, we conducted a pilot study of app effectiveness on 18 participants who used the app for 2 weeks. The participants consisted of 9 patients and 9 healthy subjects. Results of Principal Component Analysis and K-Nearest Neighbours classifier show 73% accuracy in distinguishing the patients and non-patients based on their mobile sensing data.

To overcome the overheads in transferring and processing data on centralized servers, and associated privacy risk we propose a Federated learning framework for MHMS and developed an anomaly detection algorithm that solely runs on the smartphone and does not transfer raw data to the server. The algorithm helps to detect abnormal changes in a patient's behaviour thus enabling timely intervention. This anomaly detection algorithm can serve as the foundation for on-device data labeling and training in the Federated learning setting. We conducted a feasibility test for the anomaly detection algorithm, and it proves to be efficient in terms of power consumption and storage utilization. The results show a 97% reduction in storage space and a 67% reduction in power consumption compared to the traditional application.

1.4 Contributions

The thesis made two major contributions: One, we successfully designed, developed, and deployed a mobile application called PROSIT that aims to collect more than 23 types of mobile sensor data for mental health studies. The results from the pilot study showed that PROSIT was a reliable mobile application to continuously collect mobile sensor data and transfer it to clinical servers in near-invisible mode with emphasis on privacy, security, and optimal power consumption. The initial data analysis results show significant accuracy in distinguishing patients and healthy subjects based on mobile data. Two, we propose a Federated Learning framework for the mental health monitoring system to optimize mobile sensing and reduce privacy concerns. To implement on-device training which serves as a foundation for Federated Learning we developed an anomaly detection algorithm to detect the potential onset of *Depression*. The app is also effective for remotely monitoring patients with various mental health disorders during pandemic times like COVID-19.

1.5 Overview of Thesis

This thesis contains a detailed description of all the work carried out during the design, development, and field evaluation of the PROSIT app, in a sequence of 5 chapters.

CHAPTER 1 INTRODUCTION: This chapter introduces the thesis. It states the problem and the issues surrounding the problem addressed in the thesis.

CHAPTER 2 RESEARCH BACKGROUND: This chapter contains a review of research related to this thesis. It presents a review of 73 mobile sensing research on mental health, over 9 years and classifies them into 3 sub-domains: *Detection, Intervention, and Association*. It also presents an analysis of this literature by their platforms, year, country of research, mental health issues, and their study details. We conclude the chapter with a review of work done in the field of mobile sensing applications for mental health.

CHAPTER 3 PROSIT DESIGN AND IMPLEMENTATION: This chapter describes the steps taken in the design and development of the *PROSIT app*. From the early development process, architecture, design, and development of the app.

CHAPTER 4 PROSIT EVALUATION: This chapter contains details about the online survey, pilot study of the *PROSIT app, and Anomaly detection algorithm evaluation*. It also presents the primary research question, secondary research questions, and the detailed user study process.

CHAPTER 5 STUDY RESULTS: This chapter presents detailed data analysis and results. It concludes with a discussion on the results presented.

CHAPTER 6 CONCLUSION: This chapter summarizes the entire work and presents future research directions.

CHAPTER 2 RESEARCH BACKGROUND

In this chapter, we present an analysis of 73 mobile sensing research from the literature that has been developed over 9 years (2011 – 2020), to understand the trends and gaps concerning study type, goal, features, methodology, results, and drawbacks. We conclude this chapter with a discussion on Mental Health Monitoring Systems (MHMS) using Federated learning.

2.1 Mental Health Monitoring mobile applications

One of the steps to developing an effective application is finding relevant examples. These examples would help to understand what has been successful in that past and what has not, thereby assisting in avoiding previous mistakes by researchers in the field. The relevant examples of a mental health monitoring application for the current research are "Mobile sensing for mental health" Therefore, to fully understand the research area, we reviewed existing research in the area of mobile sensing for mental health. To collect relevant papers, we searched popular databases including ACM Digital Library, IEEE, JMIR, and PubMed. We also searched Google Scholar as a secondary source for any paper we may have missed.

2.1.1 Paper Inclusion Process

We used the keywords '*mobile sensing", "personal sensing", "digital phenotyping", "smartphone sensing", "smartphone application", "mobile application" with a combination of suffixes "mental health", "mental illness", "psychiatry", "mental health tracking",* and a combination of these keywords. The search results yielded 345 papers in total. Initially, we reviewed paper titles, abstracts, and introductions of each retrieved paper, for papers focusing on the design and study of mobile sensing for mental health. In total, we were able to extract 160 papers for this purpose. We skimmed through each selected paper to eliminate papers that did not meet our inclusion criteria. Our exclusion criteria were:

- *If the paper is a review paper*
- *If the paper is not a paper about a mobile sensing app*
- *If the paper is on using smartphone sensor for mental health*

- *If the paper does not involve any study to validate the mobile sensing app*
- *If the paper is a duplicate of an already chosen paper*.

After going through all the papers, we removed 112 papers that met these exclusion criteria which left us with 48 papers. From these 48 papers, we skimmed through their reference sections for papers that also discussed the design and/or analysis of a mobile sensing app for mental health. We were able to collect an additional 92 papers and after applying the exclusion criteria again on these 92, we excluded 67 papers from them, leaving us with 25 extra papers. Therefore, in total, we reviewed 73 papers on mobile sensing applications for mental health monitoring.

We reviewed the 73 papers and coded them using the coding scheme developed and adapted from Orji and Moffat [103]. Specifically, we analyzed each paper under the following categories: Study Year, Publication venue, Type of mental health condition, Devices, Mobile application, Study Duration, Number of Participants, Type of Participant, Study setting, Age group, Tracking features, Mobile Operating system, Evaluation method, Domain experts, Analysis tools, Goal, Key Findings, Considerations for privacy/security/storage/battery, Challenges and Future directions, Country. Figure 2.1 shows the process flow for the inclusion of papers in the Mobile sensing applications literature review.

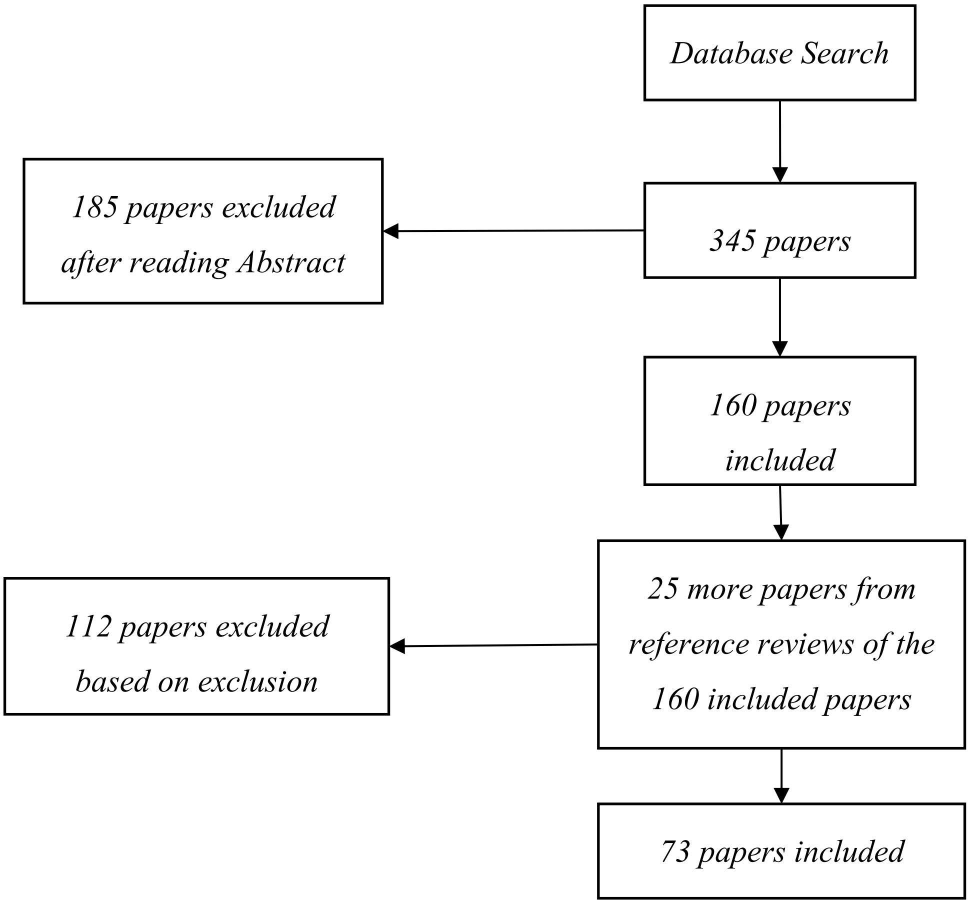


Figure 2.1 - Paper inclusion process

Figure 2.1 shows the distribution of the 73 papers across the 8 years. The first paper was published in 2012 named StressSense [77] in collaboration with researchers from Harvard University, Intel, University of Neuchâtel, and the University of Lausanne. The number of publications reached its peak in 2018 with 18 publications.

2.1.2 Targeted mental health aspects

The studies cumulatively targeted fifteen types of mental health issues, out of which *Depression, Bipolar disorder, Social anxiety, Anxiety, Schizophrenia, Borderline Personality Disorder (BPD)* and *Post Traumatic Stress Disorder (PTSD)* are mental illnesses as listed in PsychCentral [93], whereas other indicators of mental illnesses such as *Stress, Mood, Emotion, Sleep, Behaviour, Distress,* and *Loneliness* are also studied. These indicators are closely related to mental health disorders and are often found as behavioral markers to identify an underlying mental health condition.

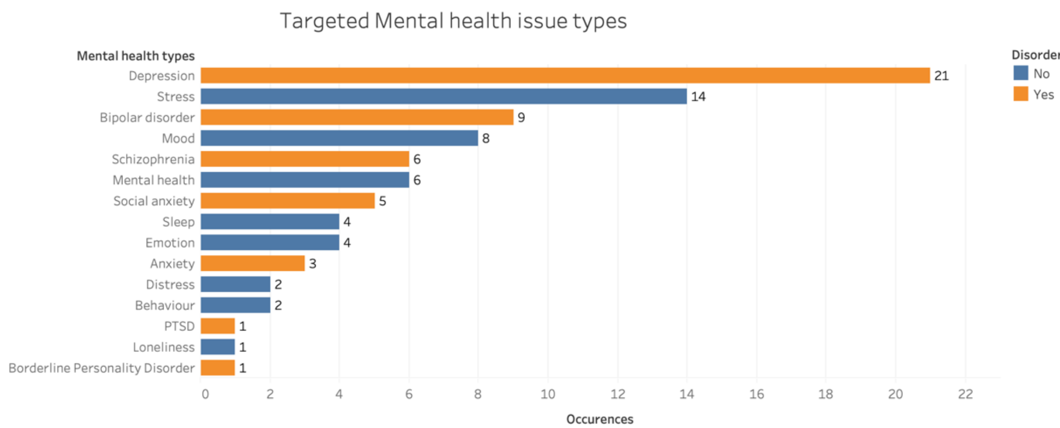


Figure.2.2 Occurrence of targeted mental health conditions and indicators

From Figure 2.2, it is evident that *Depression* is the most commonly studied mental health condition using mobile sensing with an occurrence of 21 in the reviewed literature, followed by *Stress* with 14 occurrences, and the third commonly studied disorder is *Bipolar disorder* with an occurrence value of 8.

2.1.3 Mental health illness

The descriptions of each mental health illness discussed in this paper are detailed below.

2.1.3.1 Depression

Depression is a serious medical illness that negatively affects how a person feels, thinks, and acts. It causes feelings of sadness and/or a loss of interest in activities once enjoyed. It can lead to a variety of emotional and physical problems and can decrease a person's ability to function at work and home [158].

2.1.3.2 Bipolar disorder

Bipolar disorders are brain disorders that cause changes in a person's mood, energy, and ability to function. People with bipolar disorders have extreme and intense emotional states that occur at distinct times, called mood episodes. These mood episodes are categorized as manic, hypomanic, or depressive [154]. Maniac is typically described as irritability, increased energy, and decreased sleep needs. Individuals with mania often exhibit intrusive, impulsive, and disinhibited behaviors. Psychotic symptoms can commonly occur during manic episodes. The depressive episodes are indistinguishable from unipolar depressions [169].

2.1.3.3 Social anxiety

Social anxiety disorder is when a person experience anxiety and discomfort about being embarrassed, humiliated, rejected, or looked down on in social interactions. People with this disorder will try to avoid the situation or endure it with great anxiety. Common examples are extreme fear of public speaking, meeting new people, or eating/drinking in public. The fear or anxiety causes problems with daily functioning and lasts at least six months [153].

2.1.3.4 Schizophrenia

Schizophrenia is a chronic brain disorder that affects less than one percent of the U.S. population. When schizophrenia is active, symptoms can include delusions, hallucinations, disorganized speech, trouble with thinking, and lack of motivation [160].

2.1.3.5 Anxiety

Anxiety disorders are the most common of mental disorders and affect nearly 30 percent of adults at some point in their lives. It is the anticipation of future concern, associated with muscle tension and avoidance behavior. People with anxiety disorder try to avoid situations that elicit or deteriorate their symptoms. Job performance, school work, and personal relationships can be affected [153].

2.1.3.6 Posttraumatic Stress Disorder

Posttraumatic stress disorder (PTSD) is a psychiatric disorder that may occur in people who have experienced or witnessed a traumatic event such as a natural disaster, a serious accident, a terrorist act, war/combat, or rape or who have been threatened with death, sexual violence or serious injury. People with PTSD have intense, disturbing thoughts and feelings related to their experience that last long after the traumatic event has ended. They may relive the event through flashbacks or nightmares; they may feel sadness, fear, or anger; and they may feel detached or estranged from other people. People with PTSD may avoid situations or people that remind them of the traumatic event, and they may have strong negative reactions to something as ordinary as a loud noise or an accidental touch [159].

2.1.3.7 Borderline Personality Disorder

Borderline personality disorder is a pattern of instability in personal relationships, intense emotions, poor self-image, and impulsivity. A person with the borderline personality disorder may go to great lengths to avoid being abandoned, have repeated suicide attempts, display inappropriate intense anger, or have ongoing feelings of emptiness [155]. According to DSM-IV-TR diagnostic criteria for 301.83 borderline personality disorder, the essential feature of borderline personality disorder is that it has a pervasive pattern of instability of interpersonal relationships, self-image, and affect, with notable impulsivity that begins by early adulthood and is present in various contexts, as indicated by five (or more) of the following: One. frantic efforts to avoid real or imagined abandonment. Two, a pattern of unstable and intense interpersonal relationships characterized by alternating between extremes of idealisation and devaluation Three, identity disturbance: notably and persistently unstable self-image or sense of self. Four, impulsivity in at least two areas that are potentially self-damaging (eg, spending, sex, substance misuse, reckless driving, binge eating). Five, recurrent suicidal gestures, or threats, or self-mutilating behaviour. Six, affective instability is caused by a distinct reactivity of mood (eg, intense episodic dysphoria, irritability, or anxiety usually lasting a few hours and only rarely more than a few days). Seven, chronic feelings of emptiness. Eight, inappropriate intense anger or difficulty controlling anger (eg, frequent displays of temper, constant anger, recurrent physical fights). Nine, transient, stress-related paranoid ideation or severe dissociative symptoms [170].

2.1.4 Mental health Indicators

The description of each indicator of mental health illness is detailed below

2.1.4.1 Stress

Stress can be defined as the degree to which a person feels overwhelmed or unable to cope as a result of unmanageable pressures [128]. Chronic stress disrupts nearly every system in the body.

It can suppress the immune system, upset digestive and reproductive systems, increase the risk of heart attack and stroke, and speed up the aging process. It can even rewire the brain, leaving a person more vulnerable to anxiety, depression, and other mental health problems [129].

2.1.4.2 Mood

The mood is a temporary state of mind. It is the way a person is feeling, for example, happy, sad, angry, confused, fatigued, tensed, or in simpler terms, it describes positive or negative feelings, and the corresponding intensity [9] [119]. The mood can be an indicator of mental illness like Depression and Bipolar disorder. A mood disorder may increase the risk of suicide [98].

2.1.4.3 Sleep

Sleep is essential to physical health and emotional well-being. Sleep disorders involve problems with the quality, timing, and amount of sleep, which cause problems with functioning and distress during the daytime. Difficulties in sleeping are related to both physical and emotional problems. Sleep problems can be a symptom of other mental health conditions, and they can both contribute to or exacerbate existing mental health conditions [156].

2.1.4.4 Emotion

Emotion is closely related to *Mood*, some examples of emotions or emotional states are happy, sad, relaxed, stressed, active [48] [92]. Emotional regulation is an essential feature of mental health [28], and it can be indicative of underlying mental health conditions like Stress disorder.

2.1.4.5 Distress

Distress or Psychological distress is the unpleasant feelings or emotions that a person has when feeling overwhelmed. These emotions and feelings can get in the way of daily living and affect how you react to the people around you. This distress on a higher level can lead to Attention deficit disorder, Anxiety disorder, and Depression [110].

2.1.4.6 Behavior

Behaviour is how a person acts and reacts, especially towards others. Since there is no clear biological diagnosis for psychological disorders, it is diagnosed based on observing behaviors [125].

2.1.4.7 Loneliness

Loneliness is a state of being alone and feeling sad about it. Loneliness is associated with several psychiatric disorders such as Depression, Alzheimer's disease, Stress, Personality disorder, Suicide, Sleep, Bereavement, Child abuse, and Alcoholism [99].

2.1.5 Sensing types in mobile apps

The studies analyzed used a total of 50 sensing types, with an average of 4 sensing types used in the literature. Out of which, 39 types of sensing can be captured using Smartphone alone and 11 types of sensing such as Photoplethysmography (PPG), Respiration, Skin temperature, Skin conductance, Pulse rate, Heart Rate Variability (HRV), Galvanic Skin Resistance (GSR), Cardiac activity, Ambient temperature, and Actigraphy require wearables or sensor tag to be used in conjunction with a smartphone to collect the information over Bluetooth. The sensing types are depicted in Figure 2.3

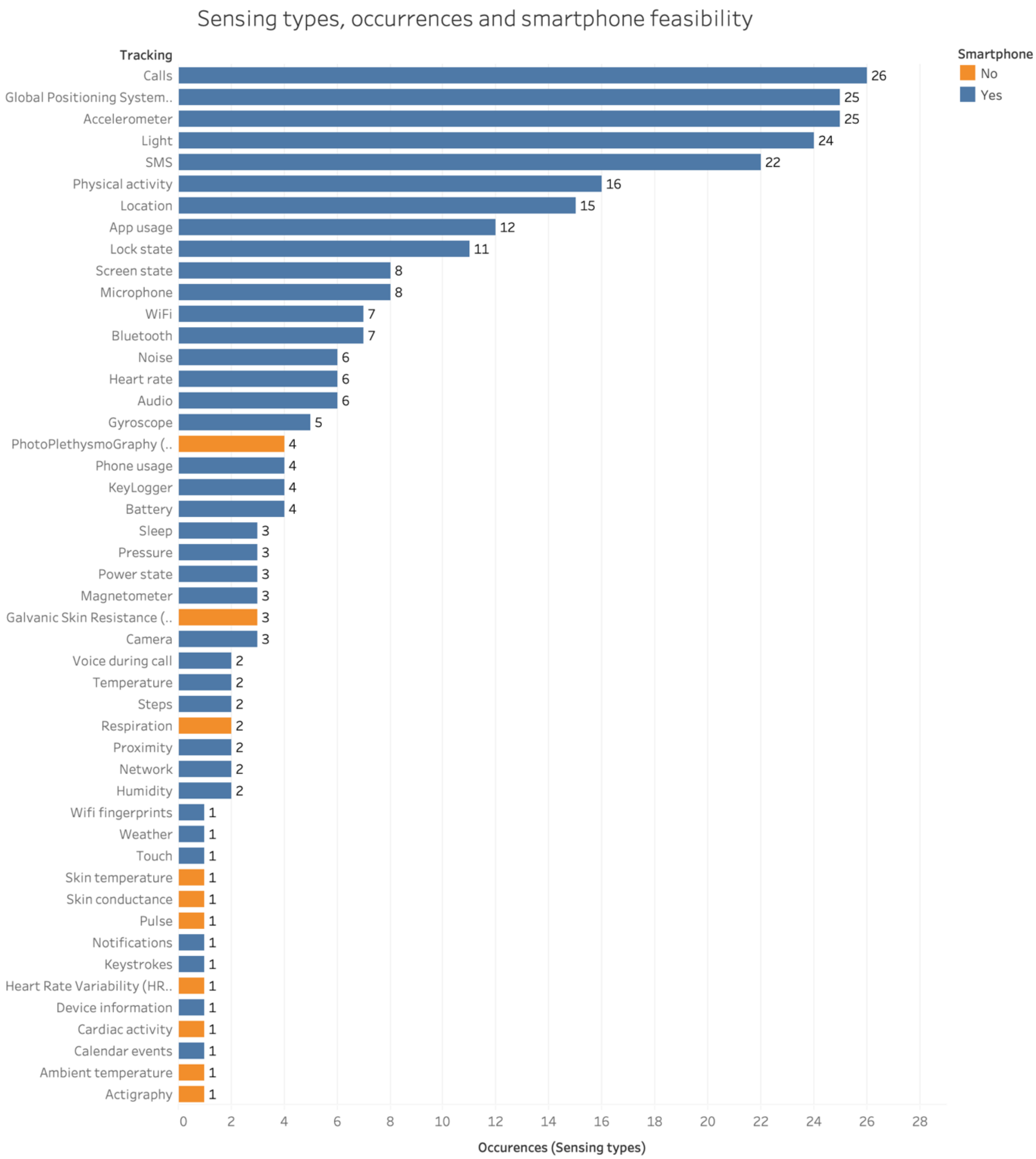


Figure.2.3 Sensing types and smartphone feasibility

Calls are the most frequently tracked feature with an occurrence of 26. Global Positioning System (GPS) and Accelerometer sensors are the most commonly used sensing type with an occurrence of 25 each, this indicates 40% of the studies utilized these two sensors. Location information is used in 15 studies, Location is different from GPS because GPS utilizes only satellite reported latitude and longitude coordinates to uniquely represent a location while

Location information can be tracked using a combination of GPS, Nearby cellular tower, and connectivity to nearby Wireless Access Points (AP). This cumulatively shows that about 55% of the studies captured the physical location of the participant to study Mental health. Other sensing types and occurrences can be inferred from Figure 2.4.

From our analysis, we classified mobile sensing apps into three major domains based on the goal of the study involving mobile sensing applications. They are Detection, Association, and Intervention.

2.1.6 Mobile sensing apps for Detection

This group of mobile sensing apps is designed to detect or predict a mental health problem using mobile sensing data and periodic self-report responses. The number of studies focused on detecting mental health issues is 35, the distribution of the studies across various mental health issues is shown in Figure 2.4

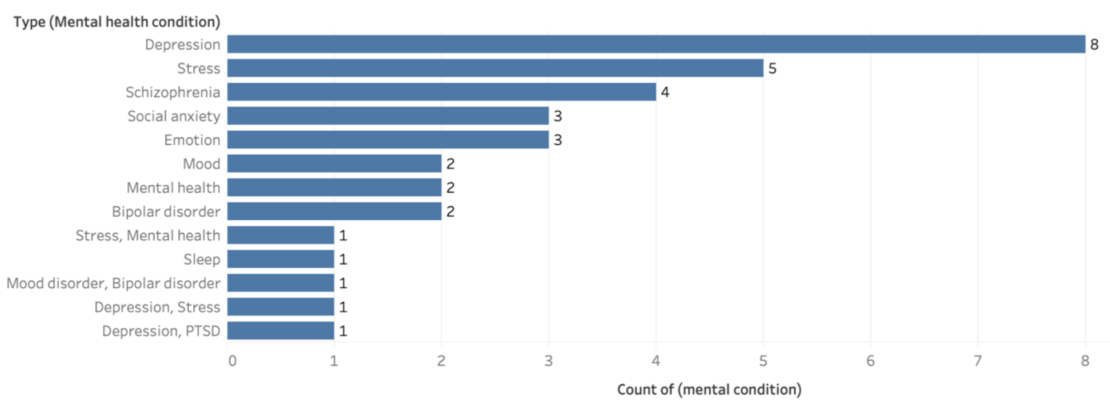


Figure 2.4 Detection - Mental condition and number of studies

2.1.6.1 Detection of Depression

Eight studies were focused on detecting Depression. Out of the seven, one study used the tool *Purple Robot* [115] to track GPS and phone usage information and found that this sensing information is strongly related to depression. One study by Farhan et al. [41] captured location, physical activity, and audio data to successfully classify participants with 87% accuracy. The *StudentLife* study [148] tracked seven types of sensing information from smartphones and wearables and succeeded in predicting weekly depression with precision 81.5% and recall 69.1%. A study used an app called *eB*^{2} [12] to study 38 patients over 3 years, utilizing eight types of sensed information from the smartphone. It succeeded in detecting changes in mobility patterns in five patients. Another study used the *SOLVD* app [25] to track seven types of sensor information from the smartphone and was able to predict depression with lower mobility and social interactions. Similarly, another study used the *Short-Term Depression Detector (STDD)* app [100] to track seven types of sensor information from smartphones and smartwatches to detect five types of clusters to visualize information. Another study by DeMasi et al. [36] used only Accelerometer data from Smartphone and Smartwatch and found irregularities in sleep and personality types as meaningful predictors for Depression.

A study by Place et al. [108] aimed to detect Depression and PTSD utilizing seven types of sensed information from a smartphone and was able to detect Depression and PTSD. One study used the *PAM* app [149] and utilized a front-facing camera of a smartphone to capture opportunistic photos detecting Depression and stress, the study reported depression can be detected with body posture of lying down. A study used *StudentLife* application to track depression in college students using a combination of smartphone and wearable sensors, and conducted a study on 83 students for 8 weeks. The results of the study state significant results with 81% precision and 69% recall [148].

2.1.6.2 Detection of Stress

Five studies were focused on *Stress* detection. The study using *StressSense* app [77] is one of the earliest studies aimed to detect mental state using mobile sensor data. The study used a combination of smartphone and smartwatch data to collect Galvanic skin resistance, Accelerometer, audio, and GPS information to study stress during interviews due to cognitive load. The study was able to personalize models for individuals with minimal data. The *Affective keys* [40] study used the Keystrokes of participants and reported a possibility to detect cues of stress from changes in keystroke pressure. A study by Yamamoto et al. [167] targeted on detecting stress levels at the workplace used ten types of sensing information from smartphones and wearable and succeeded in predicting stress with 71% accuracy. Another study by Sano et al. [117] also utilized 11 types of sensing information from smartphones and wearable, only used sensor information to predict stress levels with 88% accuracy. One study by Wang et al. [143]

captured four types of sensing information among students and reported that this method is advantageous over other baselines.

2.1.6.3 Detection of Emotion

Three studies were focused to detect *Emotion.* One study by Mankodiya et al. [82] used only photos from the front-facing camera and was able to detect different emotions. A study used the *TapsSense* app [48] and was able to predict 50% of emotion states using only the typing speed of the user. The study using the *sEmoD* app [80] used both smartphones and smartwatches to capture Galvanic Skin Resistance, Temperature, and Pulse information and succeeded in distinguishing emotional states.

2.1.6.4 Detection of Social anxiety

Three studies were targeted at detecting social anxiety using mobile sensing, one study by Xu et al. [17] on 54 college students used location and communication patterns to predict social anxiety with 85% accuracy. Another study used the *Sensus* app [64] to capture accelerometer, calls, and SMS data, this study reported that they observed difference in behavioral markers before outgoing phone calls in students with high and low social anxiety. A study used calls, location, and accelerometer sensors to detect social anxiety in college students. The results show a significant difference in location and motion between anxious and non-anxious students, as well as during calls [50].

2.1.6.5 Detection of Mental Health

Two studies used the *Socialise* [16] and *AMoSS* [107] tools respectively and studied general Mental health. Socialize study focused to uncover technical challenges and user acceptability by running an actual study collecting mobile sensing information from Bluetooth and GPS. AMoSS study collect four types of sensing data and reported personalized models are better than general models for predicting mental health.

2.1.6.6 Detection of Bipolar disorder

Three studies by Osmani [105], Grünerbl et al. [55], and Zulueta et al. [168] were focused on Bipolar disorder. All the studies used only smartphone data, were in the first study [105] they were able to detect state changes and in the second one [55] researchers were able to detect early changes in the patient. In a study using *BiAffect* [168] using a custom keyboard to detect signs of Bipolar disorder and Mood, the results signify the correlation between mood states of bipolar disorder with phone usage.

2.1.6.7 Detection of Sleep, Schizophrenia, and Mood

One study by Huang et al [63] was aimed to detect *Sleep* cycles using eight types of sensing information and succeeded in detecting irregular sleep nights without the overhead of manual labeling. One study was aimed to detect Schizophrenia using the *CrossCheck* application [146]. They studied 36 patients and reported that sensor data alone has the potential to predict symptoms. Research by Spathis et al. [124] used the Emotion sense dataset which contains Accelerometer and Microphone information of 17251 participants over 3 year period, the researchers were able to improve accuracy in predicting Mood using various machine learning techniques. The study using *MoodMiner* app [79] conducted a 30-day long study on 15 participants, and results signify 50% accuracy in predicting mood data from smartphone usage data collected from the accelerometer, call, location, light, and sound sensors. The study using *Beiwe* app [134] studied the relevance of digital phenotyping on patients with Schizophrenia, the app runs on both iOS and Android operating systems and later conducted a study using *Beiwe* application [10] to detect relapse of Schizophrenia. The researchers studied patients with Schizophrenia and reported 71% higher anomalies before 2 weeks of relapse. Another study used smartphone sensors to detect Schizophrenia [61], the researchers studied 88 participants with an equal number of patients and healthy subjects, analysed the circadian routine and weekday routine of the subjects. The results show stable social routines in healthy subjects as compared to patients.

2.1.7 Mobile sensing apps for Association

This group of mobile sensing apps is designed to associate mobile sensing data with responses to self-report questions. The number of studies in this category is 33, the distribution of Association studies is shown in Figure 2.5.

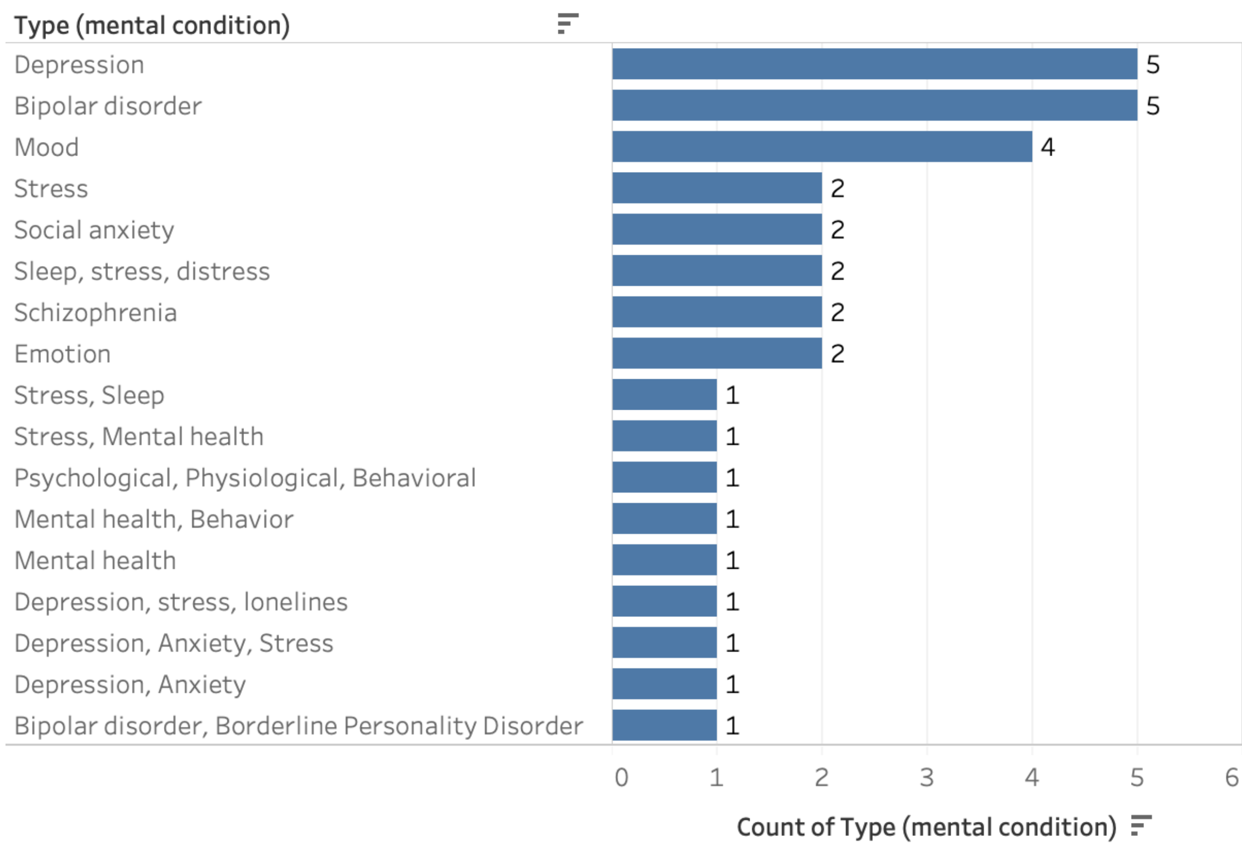


Figure 2.5 Association - Mental condition and number of studies

2.1.7.1 Association of Depression

Five studies were solely focused to identify the association between *Depression* and data captured from smartphones, out of which one study used a combination of smartphones and smartwatches for sensing. Two studies used the *Purple Robot* app [116] and *MoodTraces* app [24] to sense information from smartphones and reported that they found a strong relationship between location features and *Depression*. A study used the *MIMOSYS* tool [57] to study *Depression* in the work environment by capturing voice samples and resulted in finding a negative correlation between vitality in voice and depression questionnaire scores. One study by Mehrotra et al. [91] tracked phone usage and notifications to monitor depressive patients, the study did not find a significant relationship. The *LifeRhythm* app [78] studied 103 students for 3 months by capturing four types of sensing information using a smartphone and smartwatch and reported wearables data when combined with smartphone sensing will improve the accuracy of prediction models for *Depression*.

One study by Ben-Zeev et al. [11] aimed to study *Depression, Stress,* and *Loneliness* on Young adults using five types of sensing information from smartphones and reported that smartphones can be used as tools for unobtrusive collection of behavioral markers associated with fluctuations in mental health. Again, a study by Mestry et al. [94] captured 5 types of sensing information from the smartphone to study the association between smartphone usage and mental health during *Depression*, *Anxiety,* and *Stress,* the study was able to predict *Stress* with better accuracy. Similarly, another study by Boonstra et al. [15] attempted to associate *Depression* and *Anxiety* at the workplace with Bluetooth sensing information but did not find significant results.

2.1.7.2 Association of Bipolar disorder

Five studies were dedicated to detecting *Bipolar disorder.* One study by Osmani et al. [104] on 9 patients used five sensing features in smartphone found a correlation between physical activity and psychiatric assessment scores. The *PSYCHE* study [139] used smartphones and smart textiles to capture three types of sensing information to conclude that a personalized approach is better than a generalized approach in associating sensor data and symptoms of *Bipolar disorder*. Another study by Sabatelli et al. [114] used only Wi-Fi traces in smartphones of ten *Bipolar* patients reported a positive correlation between self-assessment scores and time spent outside the score. Another study using the PSYCHE platform by Lanata et al. [70] aimed to monitor *Bipolar* patients reported Heart Rate Variability increases with patients' clinical improvement. The *Monsenso* [30] study captured ten types of sensing information from smartphones of 129 Bipolar disorder patients for 9 months and was able to classify mood states with 89% accuracy using a personalized model.

The study using Automated Monitoring of Symptom Severity (AMoSS) study [106] recruited 100 participants with a mixed group of patients with *Bipolar disorder, Borderline Personality Disorder, and* healthy individuals with a combination of smartphone and smartwatch. They studied six types of features for four months and found regularity in data from healthy subjects, irregularities in people with the disorder, specific data from *Borderline Personality Disorder* participants who appeared more disorganized.

2.1.7.3 Association of Mood

Four studies were dedicated to studying *Mood* using mobile sensing. A large-scale study by Servia-Rodríguez et al. [119] collected data from about 18000 participants for 3 years. By collecting five types of sensing information, the study results show an ability to predict mood with 70% accuracy during weekends. This study later made the dataset publicly available as an Emotion sense dataset. The SADHealth [90] studied the effect of light variations in countries in an extreme variation on mood. The 2 yearlong study on 50 participants collected five types of mobile sensing data and reported that winter months are more depressive. Another two weeklong studies on 220 students by Majumder et al [23] collected four types of sensing information from smartphones and reported an association between negative mood and different location, activity, and time. One study by Chen et al. [27] focused on studying Mood. The participants had to take one photo every day that represents a happy moment, the results reported food as the common theme related to happy moments and social connections evoke happiness.

2.1.7.4 Association of Stress

Two studies were solely focused on associating *Stress* with mobile sensing data. Research by Ciman and Wac [29] reported capturing six types of sensing information from the smartphone in the laboratory and natural settings and was able to achieve accuracy of up to 88% in associating *Stress* levels with sensor data. One study by Park et al. [5] used PPG signals from a smartwatch along with a smartphone and reported that stress can be detected using short-term Heart Rate Variability (HRV).

Six studies were focused on associating *Stress* and other mental health indicators with mobile sensing information. Three studies studied *Sleep* along with stress. A pilot study by Tuso [136] on university students tracked 12 types of sensing information for four months and reported that they were able to see changes in sleep duration and stress levels in exam periods and breaks. A one-day study by Matthews et al. [86] using smartphone and sensor tag captured seven types of sensing information and reported that changes in light affect wellness conditions. Two closely related studies by McLeod et al. [88] and McCrae and John [87] published in 2020 studied *Stress*, *Sleep,* and *Distress* and reported ambient light and audio affect wellbeing. Two other studies by

Ben-Zeev et al. [11] and Mestry et al. [94] studied *Depression* and *Stress* is detailed in section 2.1.7.1

2.1.7.5 Association of Schizophrenia, Mental health

Two research were focused on finding an association between *Schizophrenia* and mobile sensing. A study used the *CrossCheck* sensing system [145] studied eight types of sensing data and reported a strong correlation between predicted mental health indicators and ground truth on *Schizophrenia*. A recent study by Wang et al. [151] on 55 patients with *Schizophrenia,* collected eight types of smartphone sensing information and performed extensive analysis. The study reports being able to predict social functioning in *Schizophrenia* patients accurately.

2.1.7.6 Association of Mood, Emotion, Behavior

Nine studies were focused on other factors related to mental health such as *Mood*, *Emotion*, and *Behaviour*. Two studies tried to find an association between *Emotion* and mobile sensing data. The *MyTraces* study [92] captured GPS data alone and was able to find a strong relationship between mobility patterns and emotional states. The *EmoKey* study [47] used only *Keyboard* data to determine emotional states with 78% accuracy. An early study used the *PsychLog* app [43] to collect heart rate and accelerometer data from smartphones and wearables. They reported feasibility and acceptance of the system. A 2014 *StudentLife* study [147] collected five types of sensing information from student's smartphones and found a correlation between sensing data, mental health, and academic performance. The most recent paper published in June 2020 tried to find an association between mental health and early phases of COVID-19, the study results from three types of mobile sensing data can associate sedentary lifestyle, increased stress, and depression during COVID-19 pandemic times.

2.1.8 Mobile sensing apps for Intervention

This group of mobile sensing apps is designed to deliver mental health interventions based on input from mobile sensing data and self-report questions. One research was focused on studying Mental health interventions, and three studies aimed at Detection and Intervention.

2.1.8.1 Intervention for Depression

The study using *MOSS* app [141] collected seven types of mobile sensing data to study 126 participants for 8 weeks on the effectiveness of interventions for *Depression.* The results state improvements in the participants.

2.1.8.2 Detection and Intervention for Mood, Depression, and Stress

The study using *Smartphone* app [9] collected data related to Calls, SMS in combination with Calendar events, and diaries. The app sent a notification to improve *Mood*, though no significant results were reported. A study developed a chatbot *EMMA* [46] targeting to reduce *Depression* and *Anxiety*, the app was able to detect a condition from GPS information and suggest a list of interventions to choose from. One study by Edirisooriya et al. [39] aimed to reduce *Stress* in undergraduate students, the study collected three types of sensing data from smartphones and wearable, and the result was able to detect stress with 96% accuracy and suggest relaxation techniques. Mobilyze! the study is one of the early studies in 2011, implemented on-device learning to detect depression and online intervention through the website, email, and telephone, the study resulted in an accuracy of 60% to 90% [21].

2.2 Discussion of Related Work

Based on the analysis of the 73 papers, we were able to draw some insights and conclusions on the trends and gaps in Mobile sensing applications for mental health.

2.2.1 Mobile sensing applications by operating systems

From Figure 2.6, we can see that the Android operating system is the most frequently used platform for mobile sensing app development. The majority of the applications were developed to run on Android operating systems as 75% (n=41 for Android and n=14 for Android and iOS, whereas only 22% (n=14 for Android and iOS and n=2 for iOS) focused on developing apps to run on iOS platform. 18% (n=13) studies did not specify the mobile operating system, and one study from 2013 developed mobile sensing app to support the Windows operating system. It is surprising that though 98% percent of mobile operating systems are either Android or iPhone [1], only 19% of the studies were designed to include participants using both iOS and Android operating systems. The distribution is shown in Figure 2.6.

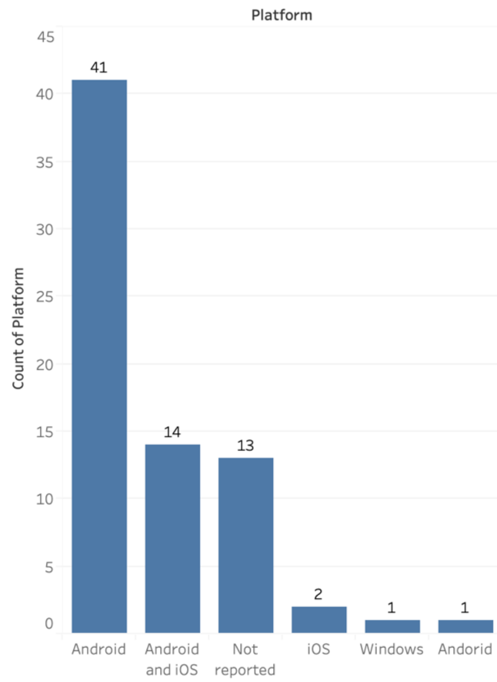


Figure 2.6 Mobile sensing applications for mental health by the operating system

2.2.2 Mobile sensing apps by country or region

The mobile sensing applications we analyzed spanned 8 years, from 2012 to 2020, and targeted a mix of various countries from different continents Figure 2.7 shows that these mobile sensing researches were carried out in 19 countries, with the USA leading with 47% of all the researches. The UK stood second with 10%, while Austria is in third place with 6% of all the apps. Only 4% of the apps emerged from Canada and Korea. Some research also took place in collaboration with multiple countries. The distribution across regions is shown in Figure 2.7

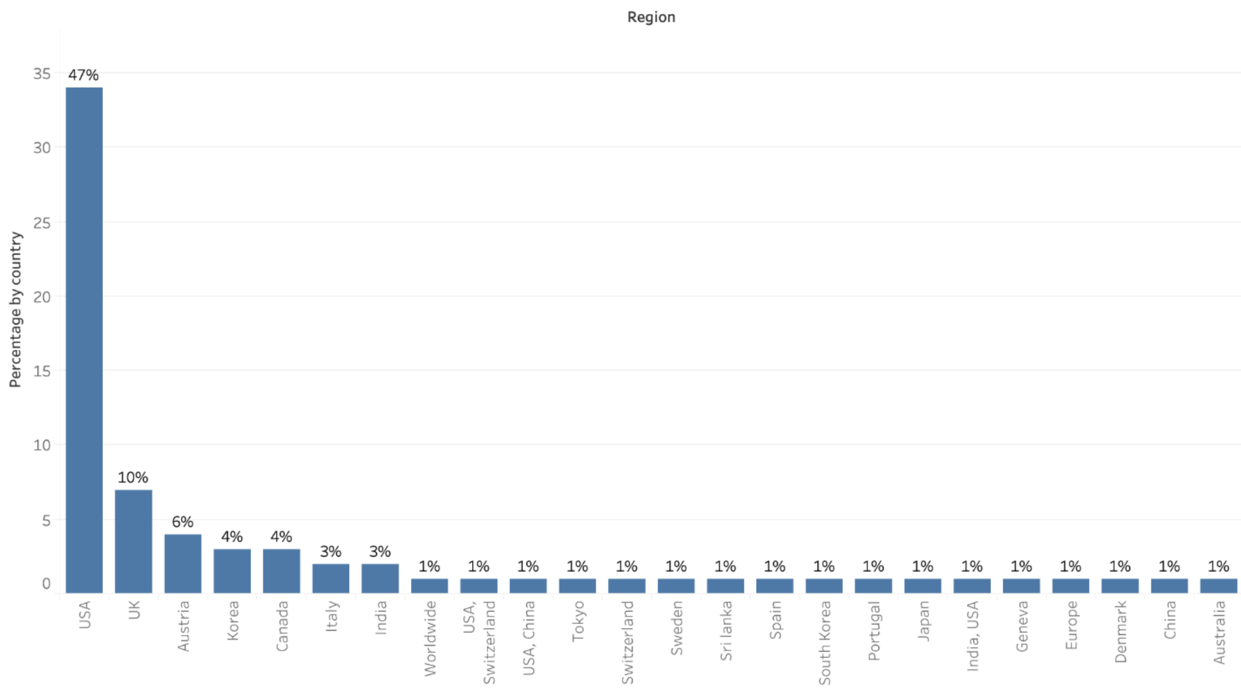


Figure 2.7 Mobile sensing application research by Country or region

Though 68% of studies emerged from countries where the number of iOS users is higher than Android users [7], countries including USA (Combining joint studies with other studies), UK, Canada, Japan, and Australia. surprisingly, only 21% of the papers studied iPhone users.

2.2.3 Mobile sensing applications for mental health evaluation details

The sample size of the mobile sensing app evaluations within the period varied significantly, ranging from 5 to 18000 participants. The study duration varied from 20 minutes to 4 years and data collection took place in five different settings laboratory, controlled environment, wild, workplace, and natural. 92% (n=59) of the data collection took place in a natural setting, which is a basic norm for mobile sensing applications, 3% (n=2) studies took place in a laboratory setting, 2% (n=1) study took place in the workplace, a combination of laboratory, wild, and controlled setting each.

2.2.3.1 Target audience

As evident in Figure 2.8, 38% of the studies were targeted at students, 21% were targeted at patients and 4% of studies focused on both patients and healthy participants in the study, 13%

were targeted at the general population mentioned as in-the-wild or wild (uncontrolled setting), while 7% study had only healthy participants. This is no surprise because most of the studies are conducted by academic institutions and 39% of studies are targeted at students or people in academia, while the actual patients with psychiatric disorders were targeted comparatively less.

2.2.3.2 Study methods

Looking at the study methods employed in the research, 81% of the studies relied on a combination of mobile sensor data and frequent questions termed as Electronic Momentary Assessment (EMA) answered by the participants, whereas only 18% of the studies used only mobile sensor data. The distribution of study methods is shown in Figure 2.9. Though regularly prompting individuals to answer questions raises challenges of participation burden [18]. This is the most used method found in the literature. This implies participant burden is not considered in 81% of the studies.

2.2.3.3 Domain expert involvement

Analyzing the domain expertise of the mobile sensing apps, 54% of the apps reported to have involved psychiatrist, psychologist, clinician, or departments related to psychiatry in their studies, and the remaining 45% of the apps were developed and studies were conducted only from the Computer science perspective and knowledge of psychiatry form literature.

2.2.3.4 Data Analysis methods

Looking at the data analysis methods, 98 unique methods were identified out of which Correlation (n=26), Regression (n=21), Support Vector Machine (n=10), Random Forest (n=6), K-Nearest Neighbors (n=4), and K-means clustering (n=4) were most commonly employed. These 6 methods cumulatively add up to 71% of total analysis methods. 95% (n=61) studies reported to have achieved significant results based on the stated research goals.

2.2.3.5 Privacy, Security, Power consumption

Though researcher evidence stated that privacy, security, power consumption should be taken into account while developing mobile sensing applications [97] [112]. As shown in Figure 2.10, surprisingly, 33% of studies did not report taking privacy, security, or power consumption into considerations in their design. Only 30% of studies considered privacy concerns of the participants and took measures accordingly, 18% of studies reported that they considered data security, 16% of studies cared about draining smartphone battery and only 1% of the studies took measures to reduce the burden on local storage. These four considerations are crucial to a mobile sensing application as mentioned in [44].

2.2.3.6 Data upload

Except for 2 (3%) of the studies, 62 (97%) studies involved uploading entire user data to a server, analyzing them, and then producing results. This produces overhead in data pre-processing, applying different analysis methods, and producing results. This aligns with the problem highlighted in this thesis which we aim to solve using the Federated Learning framework, that though smartphones are increasingly equipped with more computing power, they are simply used to collect and transmit data. Distributed computing is not employed in current studies.

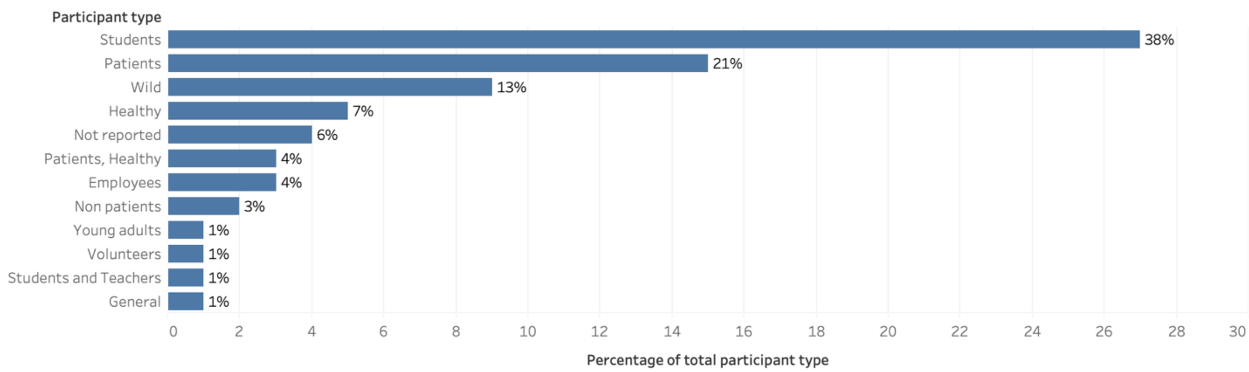


Figure 2.8 Mobile sensing apps by participant type

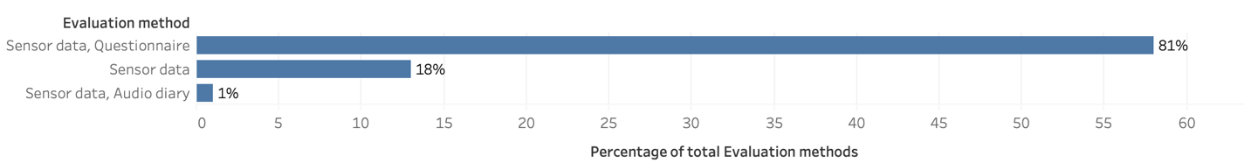


Figure 2.9 Mobile sensing apps by Study Method

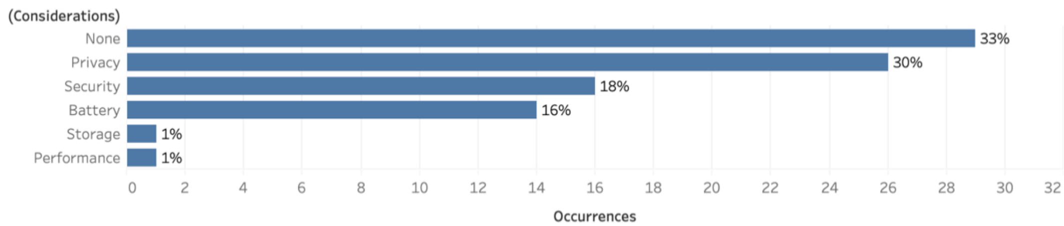


Figure 2.10 Percentage of research that took the listed measures into account

2.3 Summary of gaps in the existing literature

The gap in existing studies was identified as follows. Though more than 95% of the studies yielded positive outcomes concerning the goal of the study.

1. Domain expertise: Only 50% of studies involved people from the psychiatry domain in their design, this might be a major setback of why this method of monitoring is not adopted by clinical psychiatrists.
2. Redundant: 97% of sensing features were redundant. New applications were developed for each study, and only 7% (n=5) of applications were reused for different studies. Every application is built to track one to three types of mental health conditions, 50 unique sensing types were cumulatively used 303 times in 73 studies. This shows that the amount of research and efforts have not yet reached a mature stage to help the general public and psychiatrists to tackle mental health issues in the real world. Hence, there is a need for a reach mobile sensing application that collects a reach set of all possible mobile sensing information for studying multiple mental health conditions and comorbidities.
3. On-device analysis and learning: We could see based on the literature review that only 2% of the applications utilized the computing capabilities of smartphones. Though smartphones are equipped with computing power, they were only used as tools to collect and transfer data to the server, the collected data require extensive analysis which requires personnel with a skillset in data analysis. This is a significant overhead to clinical psychiatry.
4. Electronic Momentary Assessment: More than 82% of studies required participants to answer assessments questionnaire frequently up to 6 times a day. This can be annoying depending on the workload of the participant.
5. iOS users and Patients: Only 23% of studies focused on studying iOS users and patients each. The majority of people especially in North America use iOS [2]. Which emphasizes the need to study both iOS and Android users to avoid bias, and psychiatry patients have to be studied to help the clinicians.
6. Considerations: Only 30% of sensing applications were built considering privacy concerns. 18% of applications were reported to be secure. 16% considered optimizing power consumption and 1% applications focused on storage and performance. This

indicates that the majority of the applications were built without considering the crucial aspects essential for reliable mobile sensing tool.

2.4 Mobile sensing application for Mental health with Federated Learning

To bridge the gaps identified in the literature, we built the PROSIT app, a mobile sensing application for mental health that is built on both Android and iOS platforms, with as many tracking features found in the literature. By adding all possible tracking features, we facilitate psychiatrists to study multiple mental health conditions and comorbidities by selectively studying a different combination of the tracked data. Especially we focused on obtaining different sensing modalities from the iOS operating system, collaborated with an expert from Psychiatry, and addressing concerns related to Privacy, Security, battery consumption, storage, and performance.

Further, we developed a framework for on-device learning analysis and leading to Federated Learning, to detect early symptoms of Depression. Depression is the most widely studied mental health condition using mobile sensing, with a successful result from 19 studies found in the literature. We utilize the knowledge from literature to reduce the current issues in terms of privacy concerns, time spent in uploading entire mobile sensing data to server and performing extensive analysis at the end of the study, reduce mobile phone local storage needs, optimize battery power consumption and reduce internet data usage.

2.4.1 Federated learning overview

Federated Learning is a Machine Learning approach where smartphones collaboratively learn a shared prediction model while leaving all the training data on the host device. This decouples the need to store all the training data on a centralized server to do machine learning. This goes beyond the use of local models to make predictions on mobile devices but also brings model training to the device as well. Basic steps in Federated Learning are as follows: The smartphone downloads the current model from the server, improves it by learning from data on the phone, then summarizes the changes as a small focused update, sends this model update to the server using encrypted communication. The server updates from the clients are immediately averaged with other user updates to improve the shared model, the improved model is downloaded by the clients. This process continues at periodic intervals. All the training data remains on the device, and no individual updates are stored in the server [52].

2.4.1.1 Federated learning concept and applications

The concept of Federated Learning is proposed by Google in 2016 [69] and is widely used by Gboard – the Google keyboard for providing suggested words [59].

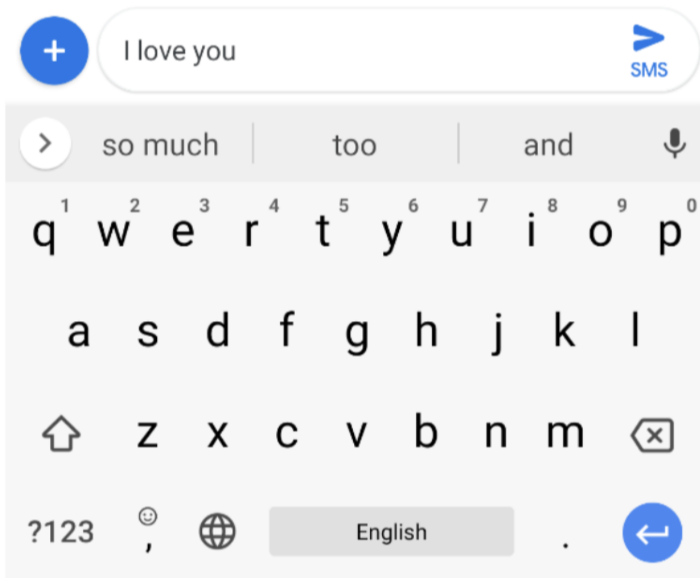


Figure 2.11 Next word predictions in Gboard. Based on the con-text "I love you", the keyboard predicts "and", "too", and "so much" [59]

Figure 2.11 shows the Google keyboard providing prediction based on the context of typed text "I love you". The suggestion strip appears just above the keyboard. It is split into three parts, the center position of the strip shows the next world with the highest probability, the second-most likely prediction is shown on the left, and the third most probable word is shown on the right side of the suggestion strip, this helps to ease the text entry.

The Federated Learning approach brings models to the data in contrast to traditional approaches of bringing data to the model. The traditional approaches require mobile devices to transfer data to a server, a large corpus of data has to be stored centrally and used to train a machine learning model, this requires continuous data logging, infrastructure, dedicated storage server, and security and privacy constraints. Even with security protocols in place, some users will be uncomfortable to share their data [19]. In the Federated Learning method, instead of uploading data from Mobile devices to the server, mobile devices locally train a model and only send an updated model to the server. Whereas the server in turn will aggregate models from all the clients to create an improved global model. This improved model is then sent to all the clients thus improving model accuracy over time without transferring data from the mobile devices.

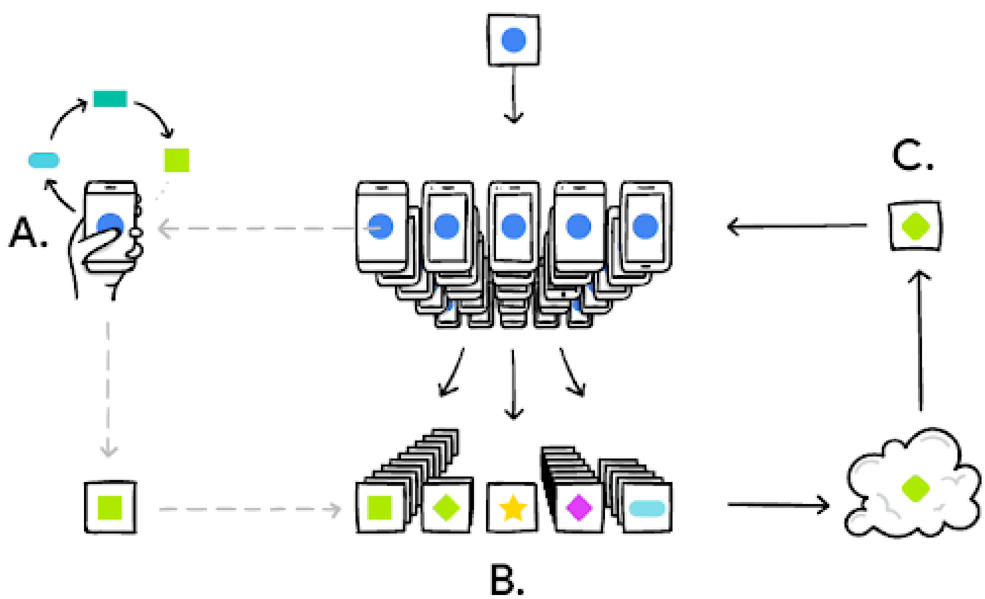


Figure 2.12 An illustration of the federated learning process from [89]*: (A) client devices compute Stochastic Gradient Descent (SGD) values and updates on locally-stored data, (B) a server aggregates the client updates to build a new global model, (C) the new model is sent back to clients, and the process is repeated*

Figure 2.12 from [89] depicts the Federated learning process in three steps. Federated learning offers security and privacy to users as data generated by the users doesn't leave the device. In traditional machine learning models, the model accuracy doesn't increase after training, but in Federated learning as the model is continuously trained and updated the model accuracy keeps improving over time. In the context of keyboard prediction, the more a person uses the Google keyboard, the next word prediction accuracy keeps improving.

Federated learning on mobile edge networks applies Mobile Edge Computing (MEC) [74] to bring Artificial Intelligence close to mobile devices where the data is produced. Mobile Edge Computing concept enables cloud computing capabilities at the edge of a network, in this case, a smartphone. Bonawitz et al., [14] address the practical issues of Federated Learning like device availability, unreliable device connectivity, device storage, and computation capabilities. The paper proposes a protocol to overcome the shortcomings by splitting the learning into three phases namely Selection, Configuration, and Reporting as depicted in Figure 2.13

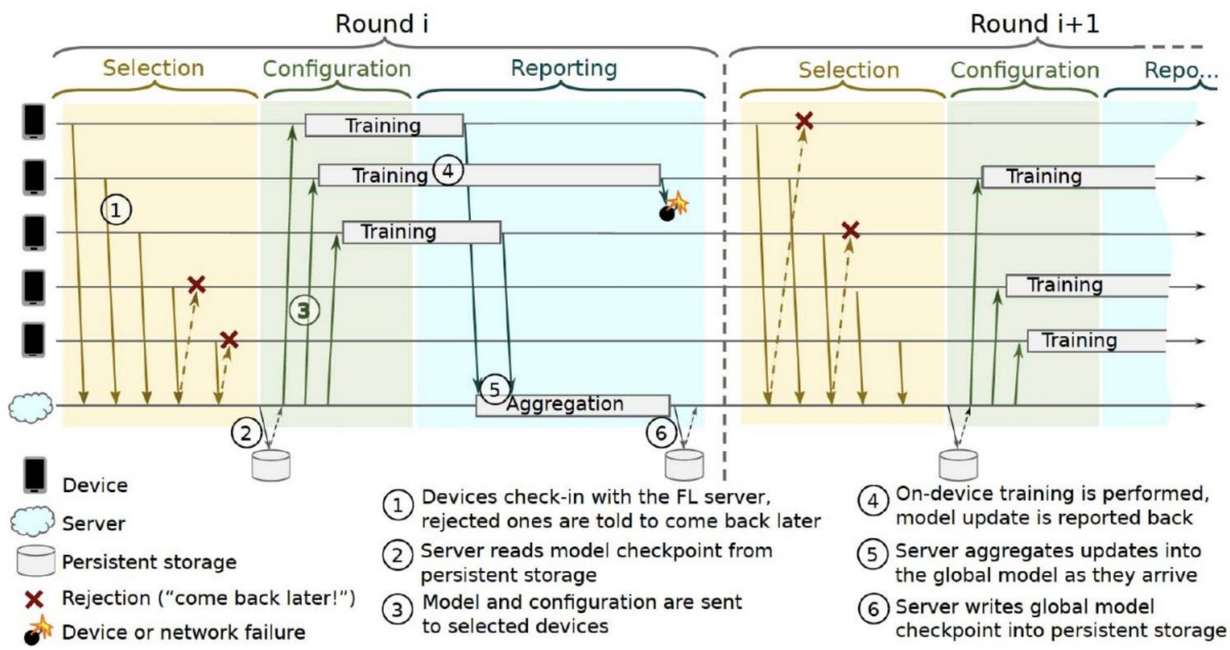


Figure 2.13 *Federated Learning Protocol* [14]

This protocol provides a solid foundation to apply Federated Learning to mobile edge devices.

2.4.1.2 Federated Learning in Healthcare

Federated Learning is being adopted in Healthcare due to the privacy-preserving nature of the method. The Model-to-data approach in Federated learning is a promising way to unlock health data currently stuck behind institutional (fire)walls and country borders, the method provides researchers access to health data without sacrificing security, confidentiality, or local control. This method can improve the quality of research outputs by providing access to more data [132].

A study by Lee et al. [73] developed a privacy-preserving platform in a Federated setting to find patient similarities across multiple institutions, the researchers were able to find similar patients across multiple hospitals with the method. This study proposes a context-specific algorithm and federated patient hashing framework.

Balsari et al [8] proposed a Federated, patient-centric Application Programming Interface (API) exchange of patient data across organizations and make informed decisions on patient data in India a densely populated country with a population of 1.3 billion. The researchers propose a framework with several levels of privacy from authorized data transfer between providers to de-identified health records for data mining for various health care entities such as doctor, hospital, chemist, specialist, payer, lab, radiology, public health agency and ministry of health.

Brisimi et al., [20] aimed to predict hospitalizations for cardiac events, solved a binary supervised classification problem where the developed algorithm performed well with relatively fewer features. The researchers conclude that the developed framework will aid in building a general decentralized optimization framework that will enable multiple data holders to converge their data to build a common predictive model without explicit data sharing.

2.4.1.3 Federated Learning for Mental Health Monitoring Systems

In the past decade, there has been significant research on remote mental health monitoring using mobile sensing techniques. Yet no research has fully addressed the privacy concern and latency in tracking sensitive participant information. With significant research and results from the domain of mental health monitoring systems, we can automate the process of the machine learning process and harness the power of mobile devices to do the prediction. Depression is the most commonly studied mental health illness (n=19), and studies suggest the significant drop in location changes and communication is an indicator of onset or ongoing depression episode. Utilizing the federated learning approach, and indicators for depression; we propose a framework for detecting depression with a federated learning approach which will eventually minimize the issues in privacy and latency thus improving remote monitoring of psychiatry patients. The development, implementation, validation of the PROSIT app and Federated Learning framework for PROSIT is discussed in the following sections.

CHAPTER 3 PROSIT DESIGN AND IMPLEMENTATION

In this chapter, we shall discuss the process that we followed before and during the design of the *PROSIT app*.

3.1 Early Development Process

To develop an effective mental health monitoring system, we have to capture indices of daily smartphone use in a natural setting. We listed all the sensing capabilities found in the literature and eliminated a few after seeking advice from psychiatrists with whom we collaborate and implemented all the possible sensing features in iOS and Android which are more likely to be predictive of mental health. The table below shows a list of tracking capabilities found in the Literature and their implementation status in PROSIT apps' Android and iOS versions.

Table 3.1 Sensing types included in PROSIT iOS and Android Versions

Sensing type found in Literature	**Description**	**PROSIT iOS**	**PROSIT Android**
Accelerometer	The accelerometer sensor measures the acceleration force in m/s2 that is applied to a device on all three physical axes (x, y, and z), including the force of gravity.	Yes	Yes
GPS	GPS uses a radio navigation system. It uses radio waves between satellite and receiver within the smartphone to provide location and time information	Yes	Yes
Light	The light sensor measures the ambient light level (illumination) in lx.	Yes	Yes
Calls	Voice call event information such as incoming, outgoing, missed, and disconnected.	Yes	Yes
SMS	SMS text information of sent and received messages using the smartphone	No	Yes
Microphone	Microphone sensor capture audio to measure noise levels in the ambient environment	Yes	Yes
Physical activity	Workout and physical activity information such as walking, running, cycling as tracked using a smartphone and smartwatch	Yes	Yes
App usage	Mobile applications installed on the smartphone, frequency, and usage duration information.	Yes	Yes
Lock state	Smartphone screen locked and unlocked states	Yes	Yes

Location	The physical location of the smartphone determined using WiFi, Cellular network, or GPS	Yes	Yes
WiFi	WiFi information to determine the physical location of the smartphone	No	No
Bluetooth	Bluetooth feature is used to periodically scan for nearby devices.	No	Yes
Screen state	Screen backlight ON or OFF state	Yes	Yes
Gyroscope	Measures a device's rate of rotation in rad/s around each of the three physical axes (x, y, and z).	Yes	Yes
Battery	Battery percentage	Yes	Yes
Camera	Front or Rear cameras to capture photographs.	No	No
Magnetometer	A Magnetometer is a sensor that measures the strength of the magnetic field around the phone from which the phone can obtain its absolute direction related to the earth's geomagnetic field [75]	Yes	Yes
Sleep	Sleep information such as duration, quality, and type. Measured directly using a smartwatch or derived from a combination of other sensor data.	Yes	Yes
Keyboard	Keyboard events, metadata on the text typed, and keystroke dynamics	Yes	Yes
Pressure	The pressure sensor measures barometric pressure in the ambiance	Yes	Yes
Voice during Call	Audio content during a voice call, and information such as pitch, tone, and intonation	No	No
Proximity	Measures the proximity of an object in cm relative to the view screen of a device. This sensor is typically used to determine whether a handset is being held up to a person's ear.	No	Yes
Power state	Power state events such as charging shut down and restart.	Yes	Yes
Humidity	The humidity sensor measures the relative ambient humidity in percent (%).	Yes	Yes
Temperature	The temperature sensor measures the temperature of the device in degrees Celsius (°C). This sensor implementation varies across devices	Yes	Yes
Network statistics	Network information related to internet connectivity such as mobile data, WiFi, and no connectivity.	Yes	Yes
Touch events	Touch screen interaction information such as tap, swipe, and scroll.	No	No
Steps	Number of steps walked	Yes	Yes
Notifications	Notification information such as the number of notifications received and application.	Yes	Yes
Wi-Fi fingerprints	WiFi fingerprints to determining user position based on strength of nearby WiFi signals	No	No
Calendar events	Calendar events to determine busy schedules	No	No

Keystroke dynamics	Keystroke dynamics is an automated method of identifying or confirming the identity of an individual based on the manner and the rhythm of typing on a keyboard.	No	No
Device Information	Device information such as smartphone model and operating system version	Yes	Yes
Storage	Device storage space information like the amount of free memory space on the smartphone.	No	No

From the literature, we listed 50 tracking features, out of which we excluded 15 features that require integration with wearable devices. Table 3.1 lists 35 tracking features found in the literature. out of which we implemented 23 features in the iOS version and 29 features in the Android version of the application. The remaining features were excluded based on the feasibility of implementation, suggestions from the psychiatrist, and restrictions in ethics approval. Tracking several features in iOS is challenging, we made adjustments to the iOS version in terms of few features as follows. Light – As tracking ambient light feature is forbidden in the iOS operating system, we continuously collected Screen brightness value, which is directly proportional to the ambient light. Microphones can be accessed in iOS only with a user-initiated action, we require the participant to manually start microphone listening during Bedtime to collect ambient noise information during sleep, whereas in Android microphone information is collected every 10 seconds throughout the day. App usage information cannot be directly tracked in the iOS operating system, so we require users to take screenshots of weekly Screen time statistics presented by iOS and add them in the PROSIT app, later information from these screenshots can be obtained by applying Optical Character Recognition (OCR), Notification is tracked similarly.

3.2 Domain Expertise – Psychiatry

Owing to the interdisciplinary nature of the application, advice from domain experts such as psychiatrists is crucial to bridge the gap between application development and adaptation in treatment. We are well aware of this problem, and we made sure the app is developed based on consistent input from the psychiatry department experts, whom we worked closely with. Therefore, the app was developed in close collaboration with a psychiatrist. The app is iteratively built based on Agile methodology [157]. The developer met with the psychiatrist once a week, to get continuous feedback. The entire development process took place over a space of one year. Psychiatrists were involved in every stage of development from requirement listing, design requirements, look and feel, setting thresholds for sensors, and regression testing., in weekly sprints. At the end of each sprint, a version of the app is delivered to the psychiatrist. The psychiatrist will install, use, and monitor the uploaded data and provide feedback. The app is updated based on the feedback from the psychiatrist.

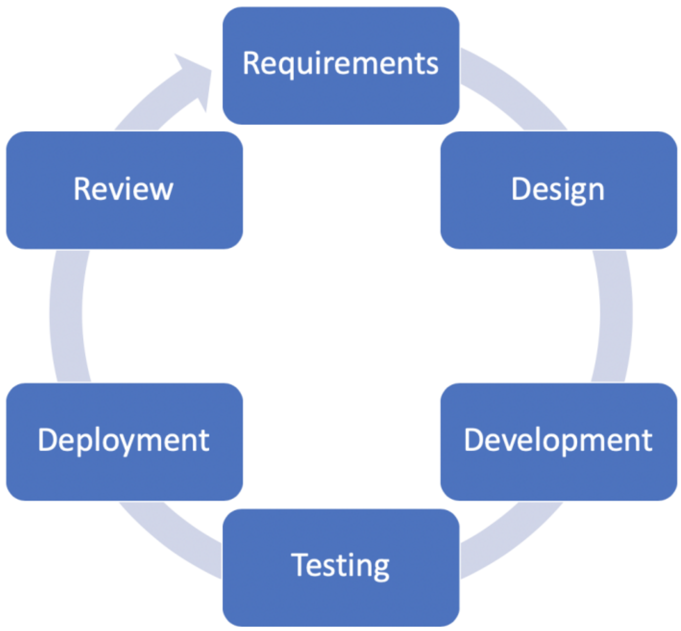


Figure 3.1 Agile development weekly sprint

The weekly sprints are discussed with the Faculty of computer science supervisor and members of the Persuasive computing lab at Dalhousie University to gather technical insights and feedbacks. Typical steps in a weekly sprint are shown in Figure 3.1. This method of development helped us to gain deep insights from different perspectives such as a psychiatrist, behavior scientists, and Computer Scientists.

3.3 PROSIT DESIGN

As indicated in the previous section, before proceeding to develop the app, we had several unstructured interviews with medical stakeholders in the Psychiatry department, to have a better understanding of the features that need to be tracked, the optimal intervals, and security measures [113] [49]. The iOS version is built-in native Swift language using XCode, Android version was built using Java, Kotlin languages using Android Studio. The iOS version design and implementation are discussed below.

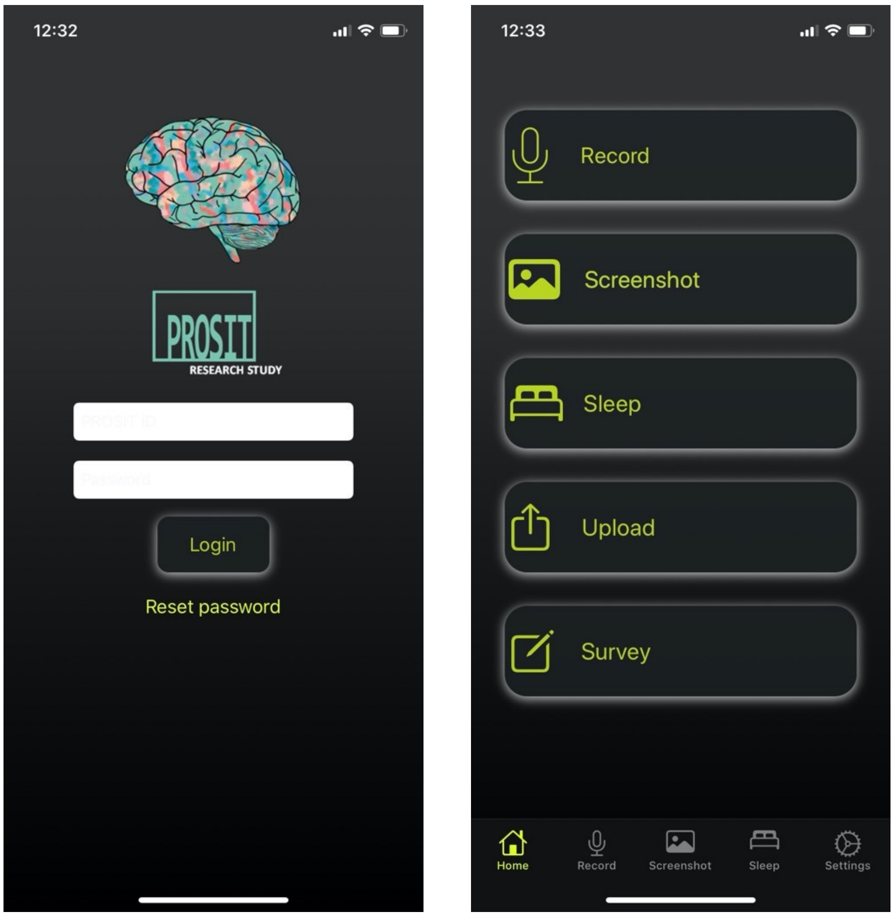


Figure 3.2 PROSIT iOS version - a) Login screen b) Home screen

Figure 3.2 shows the Login screen and Home screen of the PROSIT iOS version. Detailed screenshots of the app are provided in Appendix E

3.3.1 The Architecture

The PROSIT app has a four-tiered architecture, with two tiers on the client-side and two on the server-side. The client-side architecture lies on the smartphones and server-side architecture is hosted on the secure servers. The client and server are connected over a secure HTTPS connection. The REST API serves as a middle layer to get data from the clients and store it on the master database in the server.

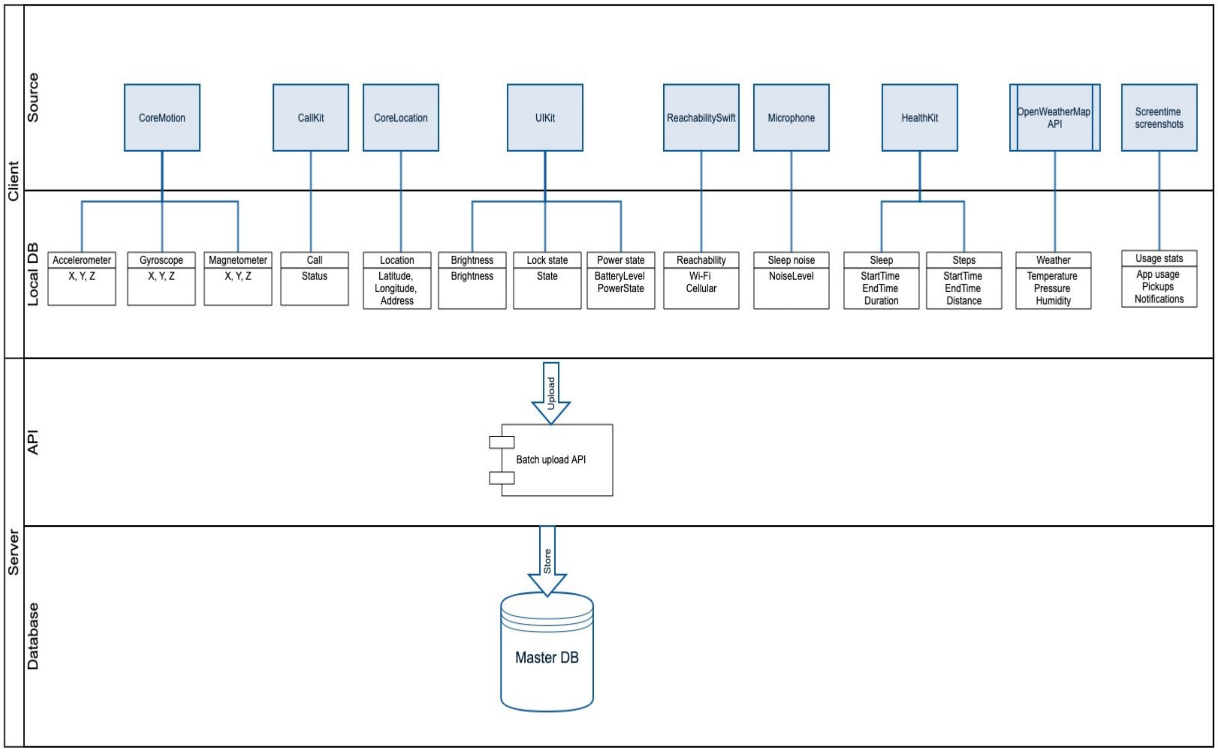


Figure 3.3 PROSIT System Architecture

Figure 3.3 depicts the four-tiered architecture split into Client and Server. The client-side comprise sources for data collection and local storage whereas the server-side comprise API and master database.

3.3.1.1 Client-side

The Client-side architecture is comprised of two layers: source and local database. The source layer lies in the iOS or Android operating system, where various types of sensing information are fetched directly from the hardware sensors or OS provided repositories, the data is periodically fetched and stored in a local SQLite database. The SQLite database serves as the second layer on the client-side. This database temporarily stores the sensing information before transmitting it to the server periodically or when internet connectivity is available.

3.3.1.1.1 Source Layer

The Source layer consists of iOS components that act as the data source to acquire sensor data. iOS developer documentation [42] lists predefined classes and repositories to obtain information from the underlying iOS operating system. The CoreMotion component is used to retrieve device motion information from Accelerometer, Gyroscope, and Magnetometer sensors in three-dimensional axis X, Y, and Z. The CallKit component is used to capture call events like CALL_CONNECTED, CALL_DISCONNECTED, CALL_INCOMING, and CALL_OUTGOING. CoreLocation component utilizes the device's location services and Global Positioning System (GPS) to retrieve location coordinates such as latitude, longitude, and Address. The UIKit component is used to retrieve device information such as screen brightness, lock states such as LOCKED and UNLOCKED, power states such as PLUGGED, UNPLUGGED, and battery percentage. The UIKit offers flexibility to set observers to capture the asynchronous events and synchronous monitoring. ReachabilitySwift library is used to retrieve device reachability status such as WiFi, Cellular, and No in terms of device internet connectivity. The microphone is directly accessed to capture noise levels during bedtime, HealthKit component is utilized to retrieve Sleep duration and Steps information automatically tracked by Apple's Health application. Similar to Apple's weather application, we use OpenWeatherMap API to retrieve weather information such as Temperature, Pressure, and Humidity. Current location coordinates are sent to this API to retrieve weather information.

3.3.1.1.2 Local database

The second layer is the local storage, this layer comprises of SQLite Database in the local storage of the smartphone, each sensing type information is stored in a separate table along with timestamp information. Information from this database is periodically sent to the PROSIT servers using the Application Programming Interface.

3.3.1.2 Server-side

The server-side architecture comprises of the Application Programming Interface (API) layer and the master database. The API built using RESTful web services serves as an interface enabling client-server communication by sending and receiving data. The master database built using MongoDB stores data sent from all the clients. This architecture is built using a combination of Nginx, Docker, and Meteor technologies for webserver, container, and dashboard respectively. The dashboard enables researchers to create participant IDs, view data, and monitor uploads.

3.3.1.2.1 Application Programming Interface (API)

The API layer on the server-side was built and maintained by the PROSIT team. This layer comprises six web services built using the RESTful framework. The RSA Key service is used to get the public RSA Key for encrypting username and password. The Token service generates a one-time token to access the token, with an expiry set to 15 minutes. This token is generated every time the user logs in to the app, changes password, and data uploads. Entry service is used to upload a single record. Batch service is used to perform batch upload and Binary service is used to upload binary files such as images and audio, and password reset service helps participants to reset their password.

3.3.1.2.2 Database

The master storage layer is the database built using MongoDB on the PROSIT server, which stores all the information sent to the server. The database has five tables namely Users, Participants, Entries, Metafiles, and Statistics. Users table stores access information about database users and administrators, Participants table holds information about the list of participants and hashed passwords. Entries table stores sensor data from the smartphones, the Metafiles table holds information about binary files uploaded from smartphones, and the Statistics table holds a summary of client upload information for monitoring purposes.

3.3.2 Design decisions for application configurations

The design decisions were made based on two factors: The features were chosen from the literature and in compliance with ethics approvals, frequency of sampling, and threshold configurations were decided based on advice from psychiatrists. We followed the Agile methodology of software development for the design and development of the PROSIT application. Every week the developer had a one-to-one meeting with the psychiatrist to discuss the design and implementation of each feature, and receive feedback on the previously implemented features if any. The design decision process for each sensing type is detailed below. A more detailed timeline on the Agile development process is detailed in Appendix J

3.3.2.1 Accelerometer, Gyroscope, and Magnetometer

The motion sensors Accelerometer, Gyroscope, and Magnetometer generate continuous data in 3-dimension at the rate of one record for every millisecond, this accounts for 86400000 x 3 number of records for every sensor per day. After discussing with the domain experts (psychiatrists), we decided to calibrate these three sensors. We live-streamed the sensor data from a test iPhone by printing the data on the XCode log while moving the smartphone around. The psychiatrists preferred to track Accelerometer to an accuracy of tap events, by setting a threshold of 0.3 on the distance between two points in 3-dimensional space. Similarly, Gyroscope and Magnetometer were set to thresholds 0.3 and 20 respectively. The calibration and threshold setting took place over multiple weeks, after using the test version of the application on different daily life activities.

3.3.2.2 Brightness

The brightness sensing feature in iPhone represents the smartphone screen brightness on a range of 0% to 100% where 0% stands for screen off state and 100% stands for full brightness. To avoid capturing duplicate values, we set up a listener that will be notified when the screen brightness changes. The data will be stored in the database only in case the current brightness value is different from the last recorded value. We discussed with the psychiatrist on setting a threshold for the tracking and decided to track all brightness levels without a threshold value.

3.3.2.3 Call

To capture call events, we set up a listener in the application, that will be notified whenever there is a call event such as CALL_DIALING, CALL_INCOMING, CALL_CONNECTED, CALL_DISCONNECTED, CALL_ON_HOLD. These are the available states offered by the Call observer delegate library of iOS. The psychiatrists wanted to track all the five statuses and hence we included all of them.

3.3.2.4 Location

Location changes are obtained in the form of latitude and longitude coordinates. We can obtain an accurate address of the location using the coordinates. In discussion with psychiatrists and considering ethics restrictions, we made multiple decisions such as removed building numbers from the address, set the precision to four decimal places, and record only if the location changed in a range of more than 20 meters. These decisions were made to preserve user privacy. The precision value of four decimal places was set after several trial and errors using Google maps, we obtained raw co-ordinates from the application and entered them in Google maps to set the precision value that does not accurately predict the current building number.

3.3.2.5 Lock state

Lock state sensing type tracks two states locked or unlocked. During the discussion with the psychiatrist, they wanted to track both the events along with the timestamps. We added a tracking setup a listener feature that gets notified whenever the lock state event changes and the data, in turn, gets stored in the database.

3.3.2.6 Power state

Power state tracking captures information about battery percentage and battery state such as charging, battery full, unplugged from the power socket, and unknown. During the discussion, the psychiatrist wanted to track all possible battery states, and we implemented listeners to capture the four states. Since the rate of battery drain in the smartphone is a gradual process, the tracking feature is decided to be set to once every four hours.

3.3.2.7 Reachability

Reachability sensing feature is implemented using ReachabilitySwift library. It offers to track changes in smartphone connectivity in three states namely WiFi, Cellular, and No connection. The psychiatrist wanted to track all three states and we implemented them using a listener that will be notified whenever there is a change in the reachability state. In addition to that, we also collect and save this information every 15 minutes to avoid malfunctioning of the third-party library used.

3.3.2.8 Sleep and Steps

Sleep and Steps data were retrieved from the HealthKit API provided by the iPhones. The data collection method and frequency are not controlled by PROSIT. Since sleep is a daily event, we decided to retrieve both sleep and steps data daily from the HealthKit and upload it to the server. This was also agreed upon by the psychiatrist.

3.3.2.9 Sleep Noise

Sleep noise captures ambient noise levels using a user's smartphone microphone. We calibrated this sensor by live-streaming the data, referred to as noise level guides which are typically referred to in 10 dB thresholds [60]. We decided to set the threshold to 5 dB to capture subtle changes within a defined noise level. Psychiatrists decided this level capture differences in snoring and capture sleep talking in patients.

3.3.2.10 Weather

Weather sensing feature using location coordinates and queries a third-party API to get current weather information of the location. Since weather changes are typically reported in daily to hourly changes, we set this threshold to 1 hour. This threshold also helps to reduce power consumption in making two external calls concerning location and weather API.

3.3.3 Client-side process flow

The High-level process flow in Figure 3.4 shows the ideal control flow structure for the PROSIT app on the client-side.

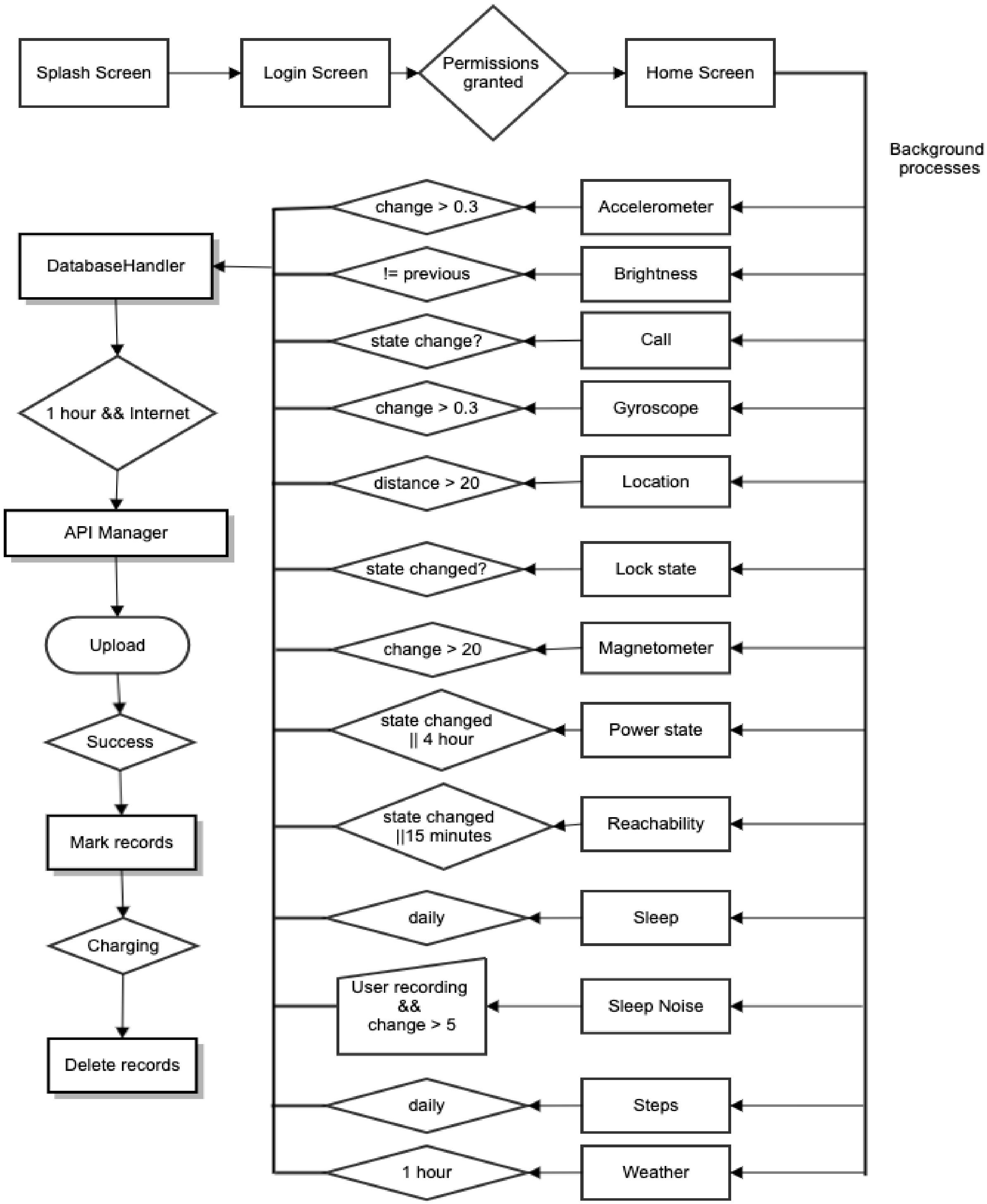


Figure 3.4 High-level overview of app process flow

When launched, the user will be presented with the splash screen and then the login screen. After a successful login, the user will be presented with several permissions that enable PROSIT to track various sensors in the background. Once permissions are granted and the user lands on the home page, all the background threads are triggered.

The user can also have manual interactions with the app in four scenarios. Sleep noise, Audio samples, Screen time, Survey. To track Sleep noise, the app will notify the user every day at 9:00 PM as a reminder to track sleep noise, the user has to open the app and start Sleep noise tracking. If there was an interruption to the microphone track, like an incoming call the tracking stops, after the call ended, the user will again be notified to resume the sleep tracking. The manual intervention is essential in sleep tracking since iOS does not permit an app to start a recording without a manual prompt from the user. Once started the sound levels will be continuously recorded. Once every two weeks, the user will be notified to record "The most exciting event of the past two weeks". The user has to record a 90-second audio recording talking about the most exciting event of the past two weeks. This feature will help clinicians to gain insight into the mental health state of the user. To capture App usage, pickups, and notifications, the user will be notified every Sunday at 5:00 PM to take screenshots of the Screen time statistics from iPhone settings and upload them in the PROSIT app. The Survey feature notifies the user to answer a seven-item survey once every three days, this feature is currently used to study the COVID-19 impact. All the above four manual inputs are uploaded to the server in the scheduled periodic uploads.

3.3.4 Privacy and Security

To preserve privacy, the following sensing features are equipped with additional conditions to avoid tracking raw data. Call information is tracked only in terms of call events, but not the contact and audio content of the call. Location coordinates are rounded to four places from 8 places, and the building number is removed from the address to capture only an approximate location of the participant. The sleep noise feature is a manual user-initiated process, which utilizes a microphone to record sound levels during bedtime. The decibel values and changes in decibel (dB) greater than 5 dB is extracted from the raw audio signals from microphones, only decibel values are stored on the database and audio information is discarded.

3.3.5 Power optimization

Mobile sensing applications have a significant impact on the battery, as the application deploys multiple tracking features running in the background. location and weather sensing features are optimized, these two sensing types consumes more battery power because location services involve external communication. Location services communicate with satellites to receive GPS coordinates and weather information is retrieved from a third-party API called Openweathermap [152]. The communication to and from the device has a significant impact on the battery. The location feature is optimized to track only if the user moves more than 20 meters and Weather information is retrieved only once an hour. Sensing types brightness, calls, lock state, power state, and reachability features are setup with listeners, that is the iOS operating system notifies these listeners only when an event related to the listener occurs. Events include screen brightness changed, incoming call, outgoing call, call connected, call disconnected, screen locked, screen unlocked, charging, unplugged, connected to Wi-Fi, connected to cellular data, disconnected from the internet. During idle times, these sensors do not run or track information. Additionally, time intervals are set to sensing types of power state to four hours, reachability to 15 minutes, and weather to one hour. These timers will trigger on set intervals, capture information, and goes back to idle mode.

3.3.6 Memory optimization

Mobile sensing applications require local storage of sensing data and transmitting to the server to cope up with the asynchronous nature of internet connectivity between client and server. To optimize memory usage, we identified sensing types that produce a huge amount of data. Accelerometer, Gyroscope, and Magnetometer are the sensors that produce records every millisecond, considering data form three axes X, Y, and Z from three sensors that account to 60000 * 3 * 3 = 540000 records in one second which will account to 46 billion data points per day from one participant. To overcome this, we have set the threshold to these three sensors, Accelerometer and Gyroscope are set with a threshold of 0.3 and the magnetometer with 20. Only when a significant change is observed from the previous value, the data gets stored in the local database. Sleep Noise is also set with a threshold of 5 dB, to only track significant changes in noise levels. Listeners discussed in 3.2.4 also account for reduced data storage. As location changes are tracked with a threshold of 20 meters, in turn, reduces the number of data points stored for the location sensing.

3.3.7 Security

To ensure privacy and security, raw username and password are not saved anywhere on the client application or the server. The client application stores RSA encrypted username and password and discards the plain text. The server stores the hash value of passwords, as Hashing is a one-way mathematical function, we cannot reverse a hashed password to its original plain text value. The login sequence is explained the Figure 3.5

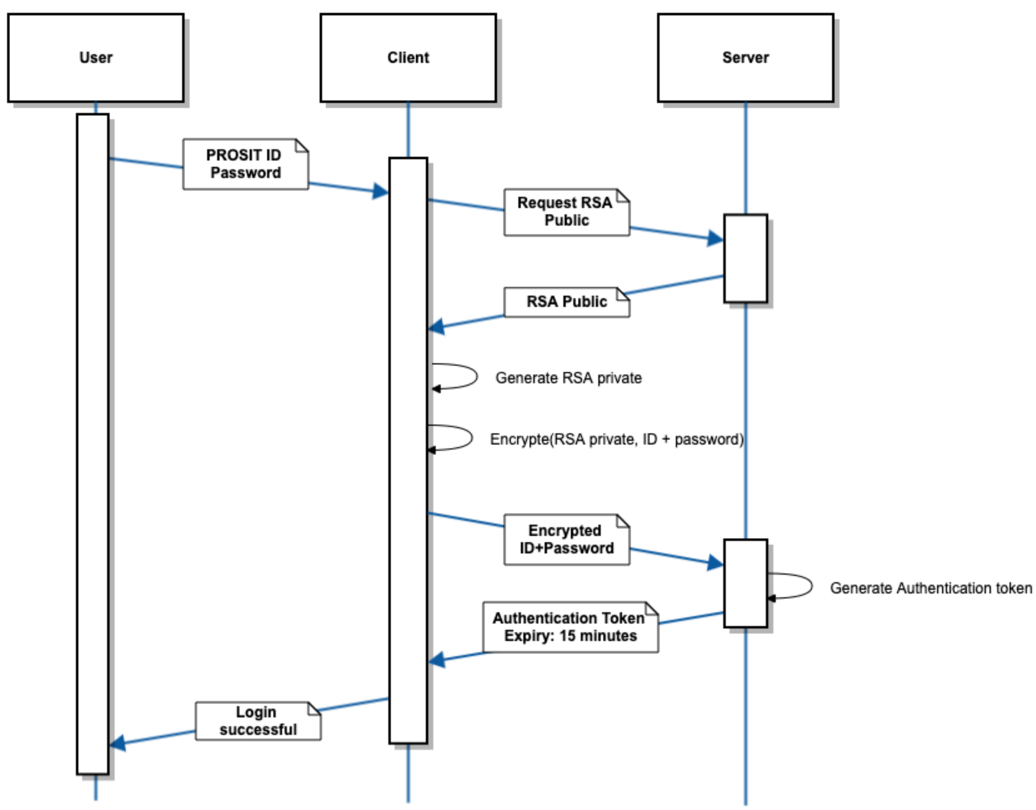


Figure 3.5 Login sequence diagram

The credentials are encrypted using ciphers RSA/NONE/OAEPWithSHA1AndMGF1Padding encryption, and the communication between client and server is over HTTP secure internet. Each client is provided with a Token valid for 15 minutes to complete login authentication and upload activities, to avoid prolonged sessions to prevent potential DDoS attacks.

The upload sequence is depicted in Figure 3.6

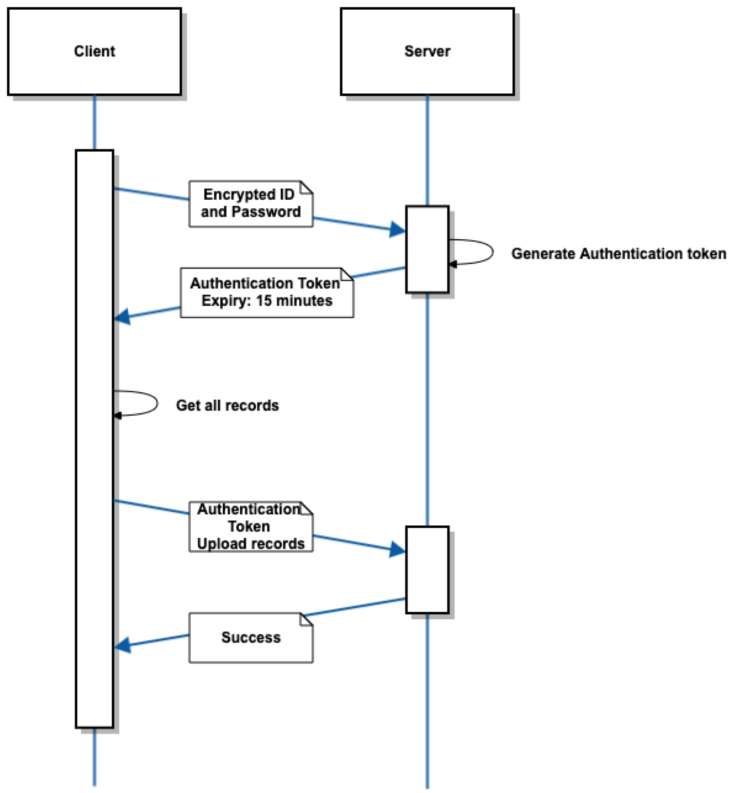


Figure 3.6 Upload sequence diagram

During an upload process, the Client sends encrypted user ID and password to the server, the server verifies the encrypted credentials sent by the client and generates an Authentication token valid for 15 minutes, the client then reads all the stored records from the local database and compiles it into JSON structure, and sends them to the server along with Authentication token in the header, after data upload is successful, the server sends back success message. After receiving a success message, the client then resumes sensor data collection, marks previously uploaded records which will then get deleted when the mobile phone is connected to a power source.

3.3.8 Federated Learning for Mental Health Monitoring Systems

To implement Federated Learning to a Mental Health Monitoring System like PROSIT, proven results from past research have to be extracted and the machine learning process should happen on the mobile device. Depression is commonly studied and can be predicted from significant changes in regular activities like going out and phone call frequency. To predict drastic changes in routines, we developed an algorithm that predicts a potential depressive episode.

To start with, the PROSIT app maintains a local database, that maintains the mean values for three epochs of a day 09:00 AM to 06:00 PM as Day, 06:00 PM to 12:00 AM as Evening, and 12:00 AM to 09:00 AM as Night as described by Wang et al., [147] in the *StudentLife* study. NumLocations stands for the number of distinct locations captured from the Location services feature of a smartphone, this value provides an insight into the outdoor movement of students and office employees who have a regular pattern of location changes for traveling to and from school and work respectively.

Since location services are energy-consuming and can be turned off in smartphone settings, we use Acceleration to supplement the location feature. Acceleration is the motion calculated from the x, y, and z-axis of the accelerometer sensor in a smartphone. This value provides an estimation of the daily movement of a person. Acceleration is calculated using the below formula

$$Acceleration = \sqrt{x^2 + y^2 + z^2}$$

CallDuration stands for the duration of calls which can be retrieved from smartphone call history as the time difference between CALL_CONNECTED as represented by TS_{cd} status and CALL_DISCONNECTED status as represented by TS_{cc}.

$$CallDuration = TS_{cd} - TS_{cc}$$

Table 3.2 Local table structure

Attribute	Epoch	Average	Type
NumLocations	Day	Local_numLocations_day	Local
NumLocations	Evening	Local_numLocations_evening	Local
NumLocations	Night	Local_numLocations_night	Local
NumLocations	Day	Global_numLocations_day	Global

NumLocations	Evening	Global_numLocations_evening	Global
NumLocations	Night	Global_numLocations_night	Global
Acceleration	Day	Local_acceleration_day	Local
Acceleration	Evening	Local_acceleration_evening	Local
Acceleration	Night	Local_acceleration_night	Local
Acceleration	Day	Global_acceleration_day	Global
Acceleration	Evening	Global_acceleration_evening	Global
Acceleration	Night	Global_acceleration_night	Global
CallDuration	Day	Local_callDuration_day	Local
CallDuration	Evening	Local_callDuration_evening	Local
CallDuration	Night	Local_callDuration_night	Local
CallDuration	Day	Global_callDuration_day	Global
CallDuration	Evening	Global_callDuration_evening	Global
CallDuration	Night	Global_callDuration_night	Global

Every day, at a scheduled time after midnight, the application will calculate the average of GForce, NumLocations and NumCalls calculated values and add/modify the local table. The structure of the Local table is shown in Table 3.2. Local data from the local table is sent to the global server, which in turn computes the global average and sends it back to all the mobile devices. The Global table structure is shown in Table 3.3. This table maintains the epoch and median information for each attribute.

Table 3.3 Global table structure

Attribute	**Epoch**	**Mean**
NumLocations	Day	numLocations_day
NumLocations	Evening	numLocations_evening
NumLocations	Night	numLocations_night
Acceleration	Day	Acceleration_day
Acceleration	Evening	Acceleration_evening
Acceleration	Night	Acceleration_night
CallDuration	Day	callDuration_day
CallDuration	Evening	callDuration_evening
CallDuration	Night	callDuration_night

Two different tables are maintained to store weekend and weekday data, which makes a total of four tables with two on the smartphone and two on the server. The total amount of records stored on the smartphone is thus 36 (18*2). Weekend and weekday can be determined based on the location-based calendar including local holidays. To maintain long-term activity patterns, six months median has to be maintained separately in addition to daily updates.

The algorithm for anomaly detection is described below

#1. Update tables

```
if internet_connected
  update_local_db
```

#2. Detect anamoly

```
for each attribute
  calculate epoch_average
    if epoch_average < 3_local_STDEV || epoch_average > 3_local_STDEV
      If epoch_average < 3_global_STDEV || epoch_average > 3_global_STDEV
        If epoch_average < 3_local_6month_STDEV || epoch_average > 3_local_6month _STDEV
          flag attribute_anomaly
          add attribute_ anomaly to anomaly _list
        else continue
     else continue
   end Calculate
end for
```

#3. Check for the severity of the anomaly

```
for each attribute _ anomaly in anomaly_list
      severity++
  else continue
end for
if severity >= 2
   trigger alert
   if anomaly == No
           update median
       else
```

```
        continue
else
  flag day

#4 Check consequent anomaly
If num_days >=30
    trigger alert
       if user_response == No
            update median
       else
           continue
end if
```

The algorithm divides each day into three epochs and keeps track of median values over a weekday, weekend, in-person values, and global values in the tables. The standard deviation from the mean is calculated to predict changes in routines, the standard deviation approach was previously used by Wang et al [150] to determine personality traits from mobile sensing data. To predict the changes to regular cycles, we calculate the Mean Absolute Deviation (MAD) introduced by Hampel [58] and if the MAD value is above 2, the value is considered an anomaly according to a study conducted to identify irregular sleeping patterns by Huang et al. [63]. When the current value is above or below two standard deviations from the mean, it is considered an anomaly and when more and more anomalies occur, the severity rating is incremented. The algorithm also accounts for anomalies across consecutive days. The trigger alert action can be supplemented by sending a notification to the user or the psychiatrist, and false alarms have to be labeled and used to train the ML algorithm. When applied over time, this approach will improve the accuracy of the model on an intra-personal level and also provide insight into global values. The global values can be distributed, and group users based on attributes such as demographic information, age, mental illness, etc.

This approach can be applied in two levels, one on the mobile device and the other on an organizational level. When applied on the mobile device, the anomaly prediction algorithm will be used to label the local dataset into NORMAL and ANOMALY, and then train the local machine learning algorithm. After training the local dataset can be discarded.

If this approach is used on an organizational level, the algorithm can be used to predict NORMAL and ANOMALY and label the dataset. Also, other variables such as participant age, comorbidity condition can be used, and then the resultant trained model has to be shared from multiple organizations to form a global model that has learned parameters from patients across different organizations.

The improved architecture of PROSIT is shown in Figure 3.7 below. For each of the three attributes, the raw values are converted into attribute weights, the average for each epoch is calculated and the list of values is fed to the algorithm. The algorithm retrieves median values from the local database and detects anomalies based on the input list. After triggering an alert, if the user acknowledges the alert, the current day dataset is labeled as ANOMALY, if the user rejects it as a false alarm, the current day's data will be labeled as NORMAL and the labeled dataset is in turn used for training the local model. After significant training for six months, the algorithm can be replaced with the local model, and the federated averaging algorithm can be applied on the global scale.

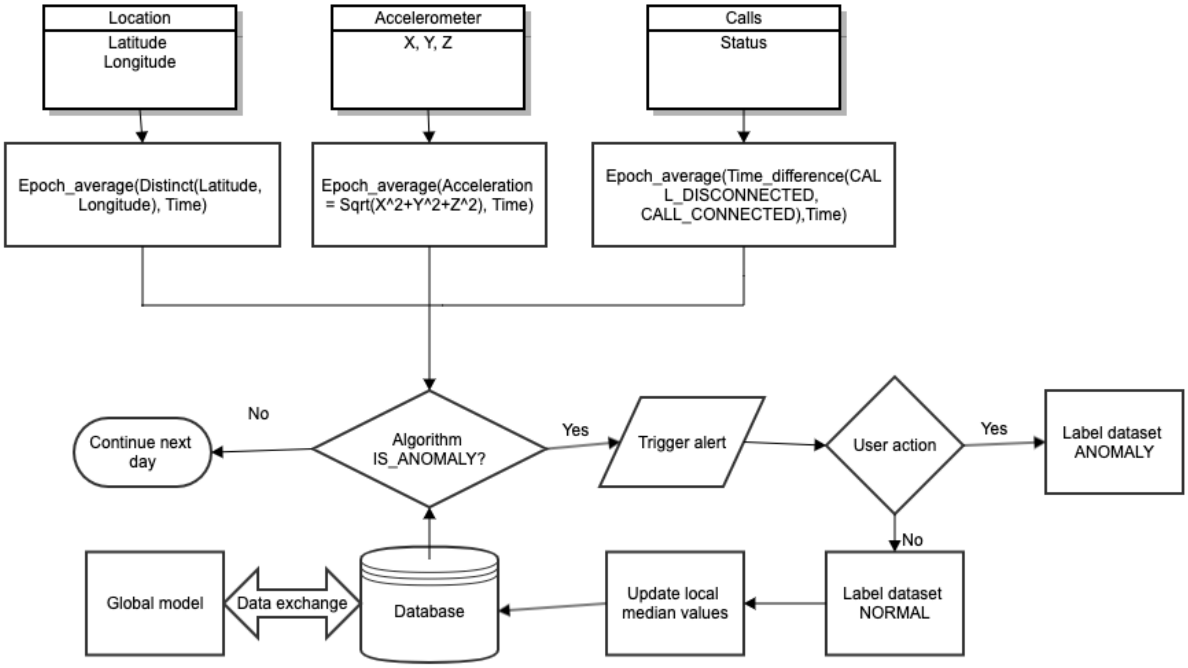


Figure 3.7 PROSIT architecture - With on-device learning

iOS applications can utilize machine learning features using the CoreML library, the latest edition of this library supports on-device training of models [31]. The library supports retraining models for Vision, Natural Language, Speech, and Sound Analysis but does not support training tabular datasets. Since the dataset of PROSIT is tabular, we currently rely only on our Anamoly detection algorithm and user action to predict anomalies in user routines. When the on-device training libraries are extended for Tabular datasets, the complete Federated Learning architecture can be implemented by utilizing on-device training.

CHAPTER 4 PROSIT EVALUATION

After developing the *PROSIT app*, we needed to evaluate three aspects of the application.

- First, we wanted to investigate public perception related to privacy concerns around tracking each type of sensor and factors that influence the privacy concerns.
- Second, we wanted to run a pilot study to ascertain that the app works well concerning being able to continuously collect and transfer mobile sensing information with minimal overhead to the user and device.
- Third, we evaluated the efficiency of the anomaly detection algorithm with respect being to detect *Depression,* the most commonly studied mental health issue from our literature review. We, therefore, developed the following research questions to guide our evaluation.

4.1 Research questions

The overarching research questions for this study are:

1. *What are the factors that affect privacy concerns related to mobile sensing data tracking?*
2. How effective is PROSIT as a mobile tracking app?
3. *Can the Anomaly detection algorithm run with optimum resource utilization?*

To be able to answer the first and second questions, we should ask the general population about their perceptions about collecting data about several indices of their daily life through smartphone sensors, and also understand the factors that influence these perceptions.

Privacy perceptions study is guided by the following specific research questions:

R1: Do people prefer data collection in the form of raw information or summarized data?

R2: What are the factors that influence privacy perceptions?

To answer the second question, we should run a pilot study by recruiting participants to install and use the PROSIT app and analyze the collected data:

R3: Can *PROSIT* effectively collect mobile sensor data and transmit it to the server continuously?

R4: Is data collected by PROSIT reliable and of acceptable quality?

To answer the fourth question, we should implement the Federated learning approach for Depression and simulate the algorithm:

R5: Does the anomaly detection algorithm help to overcome current issues related to privacy, storage, and power consumption?

Table 4.1 Shows the research question numbers and their corresponding investigations.

Table 4.1 Research Questions and their targeted outcomes

Research Questions	**Investigations**
R1	Raw data vs Summary
R2	Privacy perceptions
R3	Continuous tracking
R4	Data quality
R5	Federated learning
R6	Overcome issues

4.1.1 Approaches

To investigate the research questions, we used the following approaches. A summary of approaches is listed below, and each approach is explained in the upcoming sections

4.1.1.1 Online survey

R1 and R2 were investigated by conducting an online survey. The survey was conducted on the MTurk platform where responses from 491 participants were collected and analyzed.

4.1.1.2 PROSIT pilot study:

R2 and R3 were examined by running a pilot study. 18 participants used the app for more than 10 days and 11 million records were collected and analyzed. The process and flow of the online survey data qualification are depicted in Figure 4.1.

4.1.1.3 Federated learning simulation study

R4 and R5 were investigated by a simulation feasibility study. The anomaly detection algorithm is implemented in the iOS version, installed in a test iPhone, and 1-week statistics were collected.

4.2 Online Survey

The online survey is conducted in two steps. In the first stage, the survey is communicated to workers on Amazon Mechanical Turk (Mturk) platform by creating a Human Intelligence Task (HIT) called Mturk HIT, interested participants will accept to participate in the survey and click on the survey link provided in the HIT. In the second stage, the survey link takes the participant to the Dal opinion survey, where the participant will provide consent and answer the survey questionnaire. Upon completion of the survey, a unique completion code was generated for each participant, participants enter this completion code in the MTurk platform to indicate the person has completed the survey and can receive payment.

Completed responses are reviewed for completeness. Five attention questions were placed in random parts of the survey, If the participant has answered all the five attention questions right then the response is qualified otherwise the response is not qualified and the submission is rejected. Participants were informed about the reason for rejection reason and will not be compensated for the participation.

The steps in the online survey and qualification criteria are shown in Figure 4.1

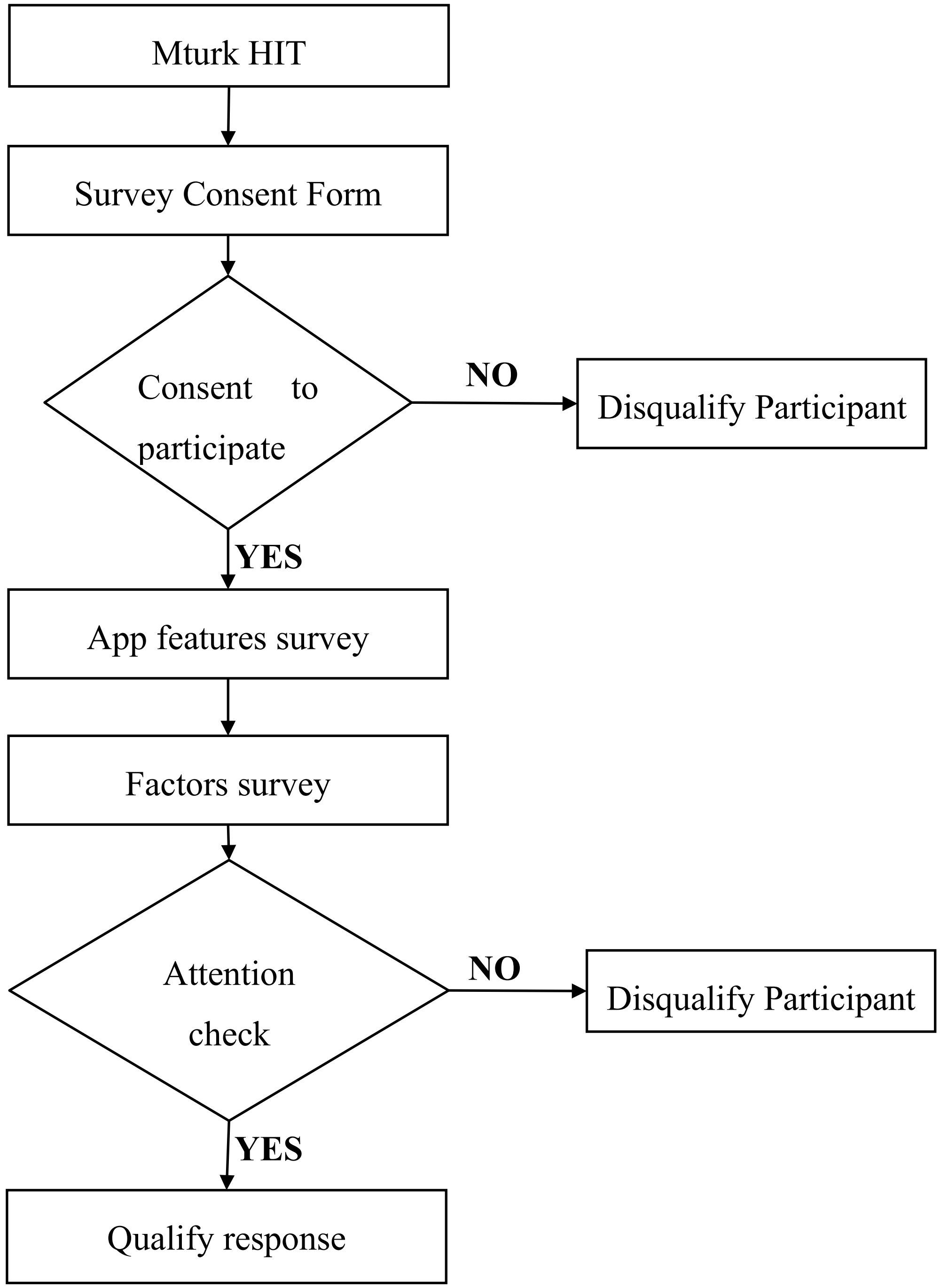


Figure 4.1 - Processes of inclusion in Online survey

4.2.1 Survey/Questionnaire Design

To answer research questions 1 to 6, we designed a study to collect data about participant's privacy perception for each sensing type, factors influencing their privacy perceptions, personality type of our participants, the impact of COVID-19 on wellbeing, and perceived usefulness of the PROSIT app during the pandemic time.

4.2.2 Answering the R1:

R1 Do people prefer data collection in the form of raw information or summarized data?

To answer this research question, we presented various sensing types implemented in PROSIT in the form of a visual representation depicting all the prominent sensing types. We presented the image shown in Figure 4.2 to the participants to provide an overview of all sensing types in the PROSIT app and asked questions in two different comparison scales. The first scale consisted of descriptions mentioning raw data collection all the time with detailed information and the second one consisted of data collection in an hourly summary. One example of the two-comparison scale is presented in Table 4.2.

This scale consists of three questions (Questions 12, 13, and 14) as specified in APPENDIX F. The questions were measured using a 5-point Likert scale adapted from [138] ranging from 1=strongly disagree to 5 strongly agree. The difference between the means of the two types of data collection (detailed and hourly summary) was determined using a Paired-Samples t-test. A greater hourly summary mean would signify a negative attitude of the participants towards mobile sensing data collection.

Figure 4.2 Depiction of sensing types in the PROSIT app

Table 4.2 Sample question on the comparison between Raw data vs Summary

Raw data	Summary
Calls information: For every call, track data about call duration, type (incoming/outgoing). Does not collect the data about phone number or conversation	Calls information: Track hourly summary of the number of calls and duration

4.2.3 Answering R2

R3 What are the factors that influence privacy perceptions?

To answer this question, we collected data about various factors that affect participants' perceptions about mobile sensing app privacy, using 13 scales adapted from Mobile Users Information Privacy Concerns Instrument (MUIPC) construct proposed by Xu et al. 2012 [164] and other scales from literature to use in our survey. Table 4.3 provides detailed information on all the factors, corresponding scales, questions along with the references.

Table 4.3 Core scales for factors affecting privacy perception

Core construct	Measurement Item	Theory
Prior privacy experience (PPE)	PPE1: How often have you personally experienced incidents whereby your personal information was used by some company or e-commerce web site without your authorization? PPE2: How much have you heard or read during the last year about the use and potential misuse of the information collected from the Internet? PPE3: How often have you personally been the victim of what you felt was an improper invasion of privacy?	Mobile users will have strong information privacy concerns once their personal information has been misused in the past or they have heard or read about the use and potential misuse of the information collected from the Internet or through mobile applications [35]
Computer anxiety (CA)	CA1: Computers are a real threat to privacy in this country. CA2: I am anxious and concerned about the pace of automation in the world. CA3: I am sometimes frustrated by increasing automation in my home.	The tendency of individuals to be uneasy, apprehensive, or fearful about current or future use of computers [126]
Perceived control (PC)	PC1: How much control do you feel you have over your personal information that has been released? PC2: How much control do you feel you have over the amount of your personal information collected by mobile apps? PC3: Overall, how much in control do you feel you have over your personal information provided to mobile apps? PC4: How much control do you feel you have over who can get access to your personal information? PC5: How much control do you feel you have	Perceived control over personal information as an individual's belief about the presence of factors that may increase or decrease the amount of control over the release and dissemination of personal information [163]. Such factors may refer to consumers' choices about the amount of information collected, e.g., through opt-in and opt-out options[26]

	over how your personal information is being used by mobile apps?	
App permission concerns (APC)	APC1: it would bother me when I am asked to accept these app permissions. APC2: I would think twice before accepting these app permissions. APC3: it would bother me to accept these app permissions.	App permission requests have a significant impact on information privacy concerns depending on the risk level and justification [56][35]
Perceived surveillance (PS)	If I would accept these app permissions... PS1: I believe that my mobile device would be monitored at least part of the time. PS2: I would be concerned that the app is collecting too much information about me. PS3: I would be concerned that the app may monitor my activities on my mobile device.	Mobile users may resist mobile apps for the fear that their activities may be watched, recorded, and transmitted to various entities. This construct is rooted in CFIP [126] IUIPC (Malhotra et al. 2004) and MUIPC [164]
Perceived intrusion (PI)	If I would accept these app permissions... PI1: I feel that as a result, others would know about me more than I am comfortable with. PI2: I believe that as a result, information about me that I consider private would be more readily available to others than I would want. PI3: I feel that as a result, information about me would be out there that, if used, would invade my privacy.	Users may resist mobile apps for the fear that the apps may interrupt their activities through unwanted presence. Intrusion can create discomfort and harm and therefore, the flow of personal information across boundaries requires users' efforts to restore their comfort levels. This is rooted in CPM Theory and MUIPC [162]
Secondary use of personal information (SU)	If I would accept these app permissions... SU1: I would be concerned that the app may use my personal information for other purposes without notifying me or getting my authorization. SU2: I would be concerned that the app may use my information for other purposes. SU3: I would be concerned that the app may share my personal information with other entities without getting my authorization.	Information is collected from individuals for one purpose but is used for other purposes without authorization from the individuals. [122]. Secondary use generates fear and uncertainty over how one's information will be used in the future, creating a sense of powerlessness and vulnerability [123]
Intention to accept app permissions (INT)	Given these app permission requests, specify the extent to which you would accept these app permissions. INT1: unwilling–willing INT2: unlikely–likely	Individuals, with higher levels of concern about information privacy practices, may be more likely in the future to refuse to participate

	INT3: not probable–probable INT4: impossible -possible	in activities that require the provision of personal information [73] [72][122]
Privacy concern (PRC)	PRC1: While using my smartphone, I am concerned about the way companies or marketers collect and use my personal information. PRC2: While using my smartphone, I feel like I am being asked to disclose a large amount of personal information. PRC3: While using my smartphone, I am concerned that a company will track me down. PRC4: It bothers me when my personal information is gathered during my smartphone use.	Mobile users who perceive that excessive personal information is collected from the smartphone have high levels of privacy concerns [66]
Trust in mobile app developers and health care providers (TR)	TR1: I think health care providers must collect my personal information through a mobile app to provide me with a better service. TR2: I can count on most health app designers to protect my data privacy.	Mobile users with higher trust in mobile application developers and health care providers (to whom the data is disclosed) have lower levels of privacy concern [66]
Privacy protection behaviour (PPB)	PPB1: Read a privacy policy before downloading a mobile app to your smartphone. PPB2: Decide not to install an app on your smartphone because you found out you would have to share personal information to use it. PPB3: Uninstall an app on your smartphone because you found out it was collecting personal information that you did not want to share. PPB4: Turn off the location tracking feature on your smartphone because you were worried about other people or companies being able to access that information. PPB5: Check the privacy setting of your smartphone.	Smartphone users who are concerned about protecting privacy by reading the privacy policy, concerned about sharing personal information, unintended data collection, and app privacy settings have higher privacy concerns [66]
Perceived benefit (PB)	PB1: Information disclosure can provide me with the convenience to instantly access the product information that I need. PB2: Information disclosure can provide me with personalized services. PB3: Information disclosure can provide me with monetary rewards. PB4: Information disclosure can provide me with entertainment.	Mobile users with a positive attitude towards information disclosure, and perceive it is beneficial, have lower levels of privacy concern[165]

Behavioral Intention (BI)	BI1: I am likely to disclose my personal information to use mobile apps for health and wellbeing in the next 12 months. BI2: I predict I would use mobile apps for health and wellbeing in the next 12 months. BI3: I intend to use mobile apps for health and wellbeing in the next 12 months.	Mobile users who intend to use mobile sensing applications and disclose personal information for health and wellbeing have lower levels of privacy concerns [164] [166]

The scale consists of 13 questions as specified in Appendix F. We measured the questions using a 5-point Likert scale adapted from [138].

4.2.4 Answering R3 and R4

R3 Can *PROSIT* effectively collect mobile sensor data and transmit it to the server continuously?

R4 Is data collected by PROSIT reliable and of acceptable quality?

To answer research questions R3 and R4, we conducted a pilot study. To answer R3, the PROSIT team recruited 18 participants to use the app for at least 2 weeks. The resultant dataset from 18 participants comprising of 9 healthy subjects and 9 patients with masked participant information was obtained from the PROSIT research team to analyze the effectiveness of PROSIT for collecting and uploading required data. Participant demographics and information about disorders were not shared to maintain the confidentiality of medical records. To answer R4 we analyzed the dataset to ensure quality and correlations are by the ground truth and performed machine learning classification to study differences in the study population.

4.2.5 Answering R5

R5 Does the anomaly detection algorithm help to overcome current issues related to privacy, storage, and power consumption?

To validate research questions R5 we did a feasibility study, we developed an anomaly detection algorithm and implemented it in the PROSIT app, then we simulated and tested the efficiency of the algorithm. We measured power consumption, memory overhead, and performance. The result provides insight into the feasibility of applying the algorithm in mobile sensing for Depression to overcome the privacy, security, memory overhead, and power consumption.

4.3 Online Survey

The online survey is conducted to understand the mobile sensing data tracking privacy preferences, factors that influence the privacy, and the impact of personality type on privacy preferences. In this section, we explain in detail the tools used and participant demographics in our study conducted as an online survey.

4.3.1 Tools used

Below is a list of all the tools and the purpose of each tool used in the online survey

- Online survey form created using Dal opinio to create the questionnaire and collect response
- Amazon Mechanical Turk (Mturk) website: To post tasks, that respondents can pick and complete
- Dalopinio server: To host the study results, and generate preliminary reports
- Amazon Mechanical Turk (Mturk) worker ID: To uniquely identify each participant in the Mturk platform reports
- Unique survey completion code: To uniquely identify each participant in the Dal opinion survey
- Participants' laptop or phones: To answer the survey questionnaire
- IBM SPSS Statistics: To analyze the results
- Tableau desktop: To create visual representations
- Microsoft Excel: To analyze results and create visual representations

4.3.2 Participants demographics Mturk Study

We created MTurk HIT requesting for participation in our study to about 1000 prospective participants across the world aged 18 and above. We excluded the data of the participants who did not answer all the attention check questions right and incomplete responses. After cleaning, we retained 491 responses.

For the included participants, we had 63% (n = 310) males and 36% (n = 179) females and less than 1% (n = 2) other gender (Figure 4.3). By country, 50% were from the United States of America, 37% were from India, and the remaining 13% were from 21 other countries (Figure 4.4). The largest age groups in our sample were the ’25-34’ group with 50%, followed by ‘18-24’ with 19%, followed by ‘45 and above’ with 16%. The smallest age group we had was ‘35-44’ which is 15% (Figure 4.5). According to the United Nations Department of Economic and Social Affairs (UNDESA), young adults fall around the age of 15 years to 35 years old [137]. Ideally, we have 69% of participants as young adults.

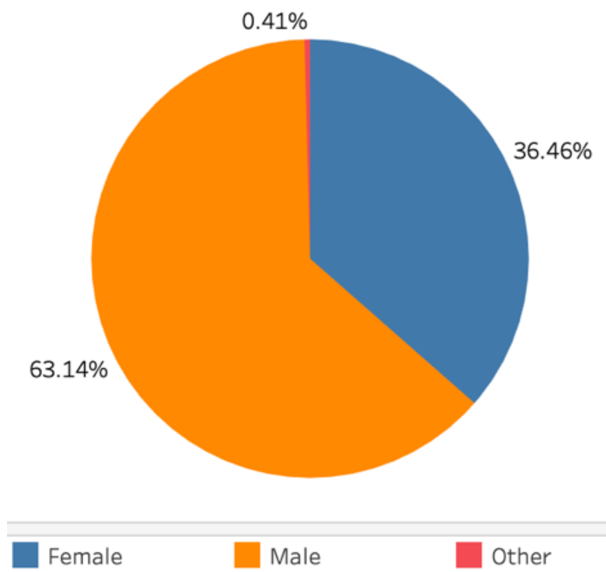


Figure 4.3 Demographics by gender

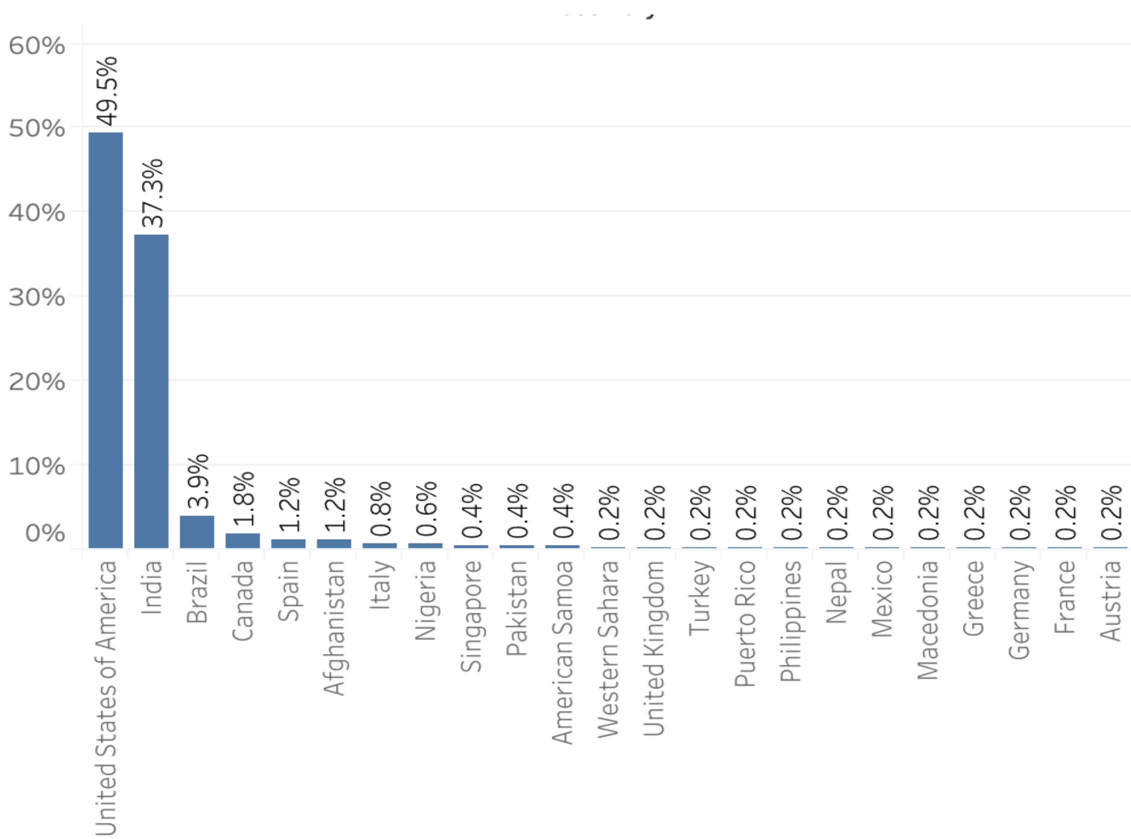


Figure 4.4 Demographics by country

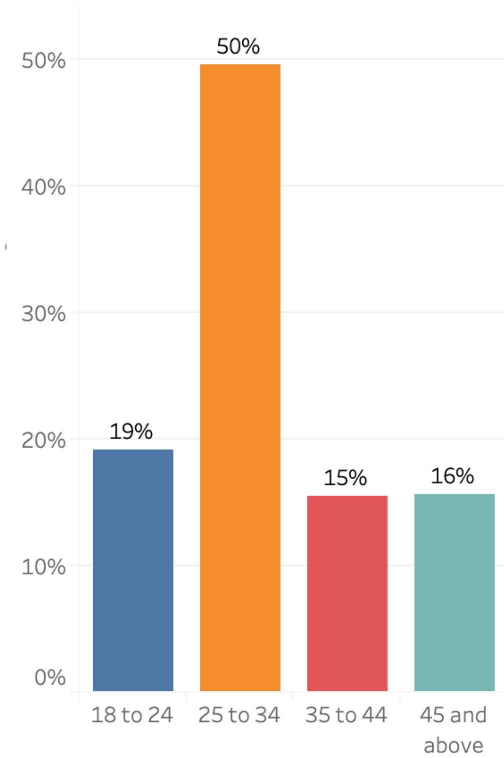


Figure 4.5 Demographics by age

By the participants' level of education, Bachelor's degree had the most 65%, followed by a Master's degree with 17%, High school or equivalent with 11%. College diploma at 5%, Doctorate, less than high school had the lowest with 1% each (Figure 4.6).

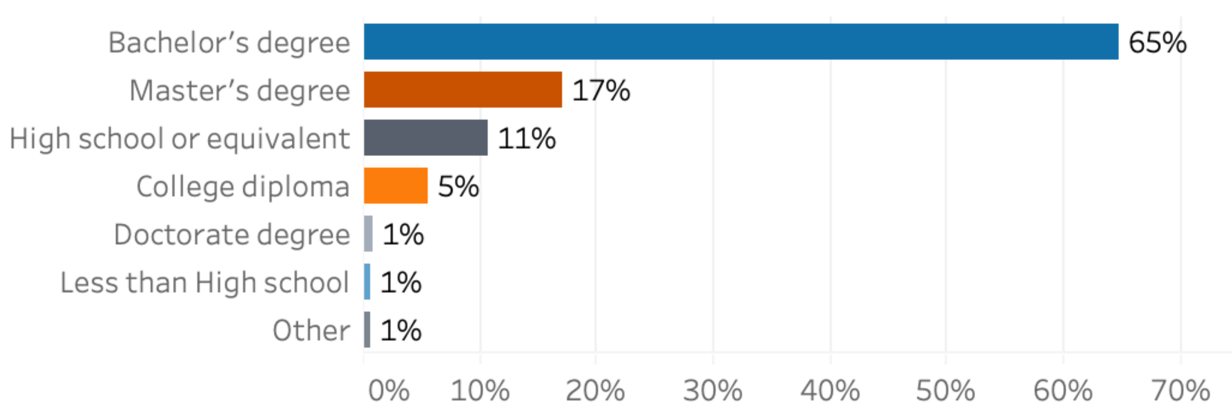


Figure 4.6 Demographics by education

50% of our participants reported that they have used a mobile app related to mental health before (Figure 4.8) and 31% of participants were diagnosed with mental illness in the past. Also, 84% of the participants reported that they have used a mobile application for Fitness. As the majority of Fitness apps involve tracking data like GPS and Activity [68]. This is crucial because the Mental Health Monitoring application has to track features similar to Fitness apps in terms of motion sensing and GPS tracking.

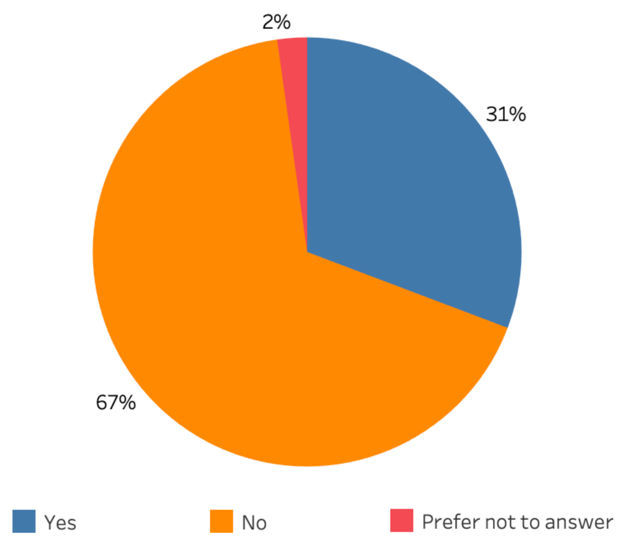


Figure 4.7 Demographics by the history of mental illness

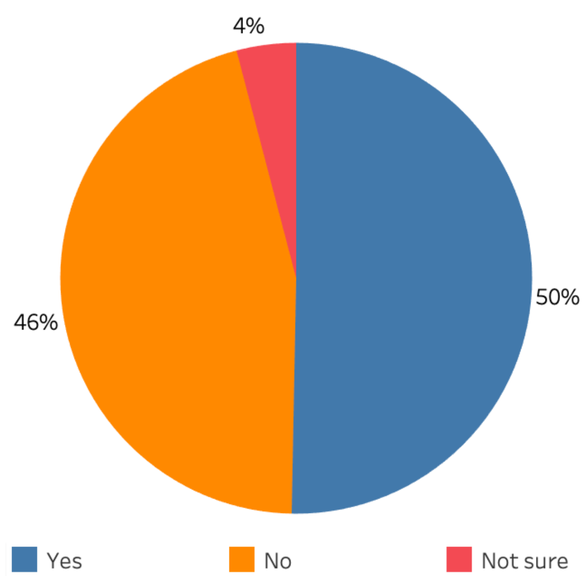


Figure 4.8 Demographics by past mobile app usage for mental health

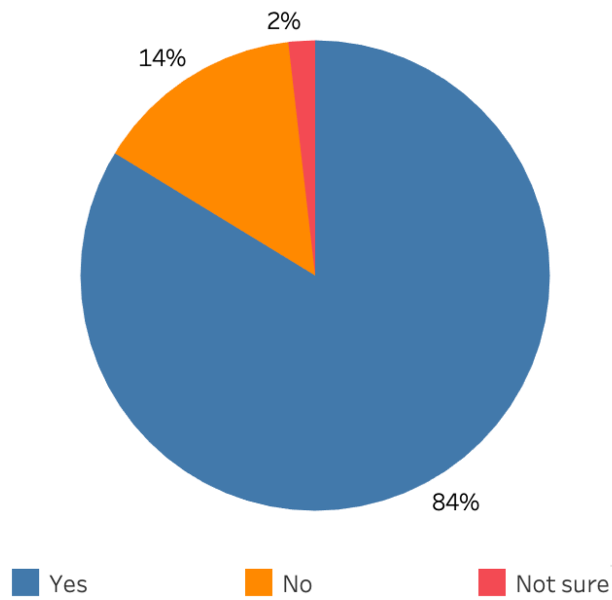


Figure 4.9 Demographics by fitness app use

99% of our participants use a smartphone, only 2 participants reported that they don't use a mobile phone. The categories based on the type of device are as follows, 72% of participants use an Android phone, 25% use an iPhone, 2% use a Windows phone and 1% uses other types (Figure 4.10).

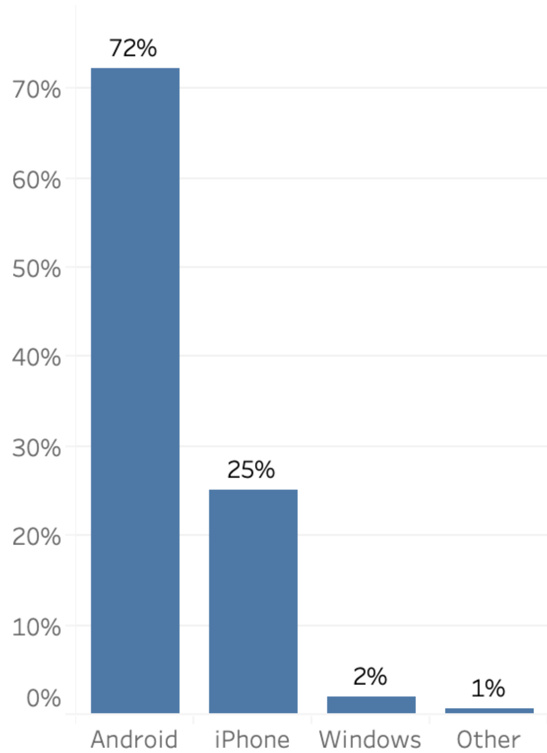


Figure 4.10 Demographics by mobile device

A summary of the participant demographics is provided in Table 4.4

Table 4.4 Participant demographics

Total Participants = 491	
Gender	Males (63%), Females (36%), Other (0.4%)
Age	‘18-24’ (19%), 25-34%’ (50%), ’35-44’ (15%), ’45 and above’ (16%)
Education	Bachelor’s degree (65%), Master’s degree (17%), High school or equivalent (11%), College diploma (5%), Doctorate (1%), Less than High school (1%) and other (1%)
Past Mental health illness	Yes (31%), No (67%), Prefer not to answer (2%)
Used mental health app before	Yes (50%), No (46%), Not sure (4%)
Used fitness app before	Yes (84%), No (14%), Not sure (2%)
Mobile phone	Android (72%), iPhone (25%), Windows (2%), Other (1%)

4.4 Pilot Study for App Functionality Evaluation

The app was evaluated by the PROSIT project team in a pilot study. The team recruited 200 participants aged 15-21, 150 participants installed the mobile app on their smartphones, and the app continuously collected data for more than 10 days from 115 participants. As part of an ongoing study, the app is actively collecting data from study participants since November 2019.

We requested permissions from the PROSIT team and received a part of the dataset (18 participants), which consisted of 11 million records. This dataset will be mentioned as a PROSIT dataset in the rest of the thesis. The 18 participants were all iOS users, located within Nova Scotia, Canada, and had data for at least 10 days. We also received data labels for the participants to group the dataset into two equal subgroups of 9 Healthy subjects and 9 Patients. The dataset contained data for participants who completed at least 10 days on the data collection before 15-March-2020. The end date is chosen because on 15-March-2020 lockdown was enforced in Nova Scotia, which resulted in dramatic changes in the lifestyle of participants in terms of reduced outdoor walking, minimal or no location changes, and change in patterns of mobile usage. To avoid bias due to lockdown and pandemic, we shortlisted only data before 15-March-2020.

4.5 Anomaly detection feasibility Study

The PROSIT app was integrated with the anomaly detection algorithm and we ran a feasibility study by installing the app on an iPhone 11 for 1 week. The version with the algorithm is named PROSITLite and aspects of power consumption, memory consumption, and network usage was measured against the standard version of the app.

4.5.1 Client-side anomaly detection algorithm implementation

To implement the anomaly detection algorithm, we modified the architecture of the PROSIT app by removing the local database that offers persistent storage for the raw sensor data and replaced it with lists that temporarily stores the sensor data, the data in lists are then used to calculate median after each epoch completes. This new version is named PROSITLite in the following sections. A daily scheduler executes at 12:05 AM nightly and executes the algorithm to detect an anomaly. A notification is set to be sent to the user at 09:00 AM if an anomaly is detected, and the user can respond with Yes or No.

4.5.2 Performance evaluation of PROSIT and PROSITLite

To test the performance of the algorithm, the versions PROSIT and PROSITLite were installed in three iOS devices running on the iOS 13 operating system for 3 weeks. Every day at 6:00 PM, data about the percentage of battery consumption for PROSITLite is collected and noted. The battery consumption information of each app is listed under the iPhone Settings -> battery. Memory information is obtained from Settings -> General -> iPhone Storage. For comparison the same procedure is followed for the regular version of PROSIT referred to as PROSIT in the following sections is the initial version of the app without the federated learning. To exclude the latencies in server communication, both PROSIT and PROSITLite were modified to perform only client-side activities of collecting mobile sensing data.

4.5.3 Performance evaluation data collection

To collect data about the memory and battery consumption of the two PROSIT versions, installed both the app versions on three devices and collected data for 3 weeks. The phones were used under natural circumstances for making calls, taken when going on walks. The data collection took place between 06-April-2020 to 19-April 2020 and 02-Dec-2020 to 14-Dec-2020. At the end of the study, the data collected is consolidated and analyzed.

CHAPTER 5 STUDY RESULTS

In the chapter, we present the results from the three studies: Online survey, Pilot study of the mobile app, and Anamoly detection algorithm feasibility study.

5.1 Online survey

In this section, we discuss the results of the Mturk study of *PROSIT app* features' privacy perception and the implications of these results. Specifically, in the following subsections, we present the results of the *privacy perceptions, factors affecting privacy perception, Impact of covid-19.*

5.1.1 Tracking Fine-grained Versus Summary Tracking

To understand the difference in preferences to track mobile sensing data in terms of tracking fine-grained user mobile activity data all the time versus collecting only summarized information in the hourly summary. Participants rated each of the 15 features on two different scales of Fine-grained and Summary.

Table 5.1 Descriptive Statistics for the two modes of tracking

	Mean	N	Std. Deviation	Std. Error Mean
Fine-grained	3.22	491	.82	.037
Summary	3.24	491	.83	.038

Table 5.2 Paired sample T-test for Fine-grained and Summary tracking

Paired Samples Test Paired differences								
	Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference		t	df	Sig. (2-tailed)
				Lower	Upper			
Fine grained - Summary	.029	.324	.015	.001	.058	2.037	490	.042

Based on the results of paired-sample statistics, we could see the differences in means are significant as $p < 0.05$.

5.1.2 Feature ratings

To get user's comfortability levels for each feature in terms of tracking data Fine-grained (FG) and in Summary, we first posted a question as "*How comfortable you feel when data from your mobile phone is collected to predict your health issues. Answer the questions assuming a mobile application has the capabilities to collect data as explained below. Rate from 1 to 5 to state how much you are comfortable in the data collection feature, where 1 being lowest and 5 being highest*", The rating for each of 15 different sensing types is detailed below. The questions were measured using a 5-point Likert scale ranging from 1=strongly disagree to 5=strongly agree.

5.1.2.1 Calls

The participants were asked to rate on two different types of call tracking with the following questions, on a comparison table.

Table 5.3 Feature rating questions for tracking All the time versus Summary

Feature	Fine-Grained	Summary
Calls	For every call, track data about call duration, type (incoming/outgoing). Does not collect the data about phone number or conversation	Track hourly summary of the number of calls and duration
SMS	For every text message, track data about length, type (sent/received). Does not collect data on the content of messages	Track hourly count of the number of SMS sent/received, and count of typed keys
Bluetooth	Track data about names of connected and nearby devices. Does not collect data on content sent/received	Track hourly count of the number of BlueTooth devices nearby
Location	Track data about every street you visit. Does not collect data on the building address	Track hourly summary of location changes in distance
Steps	Track data about the number of meters walked/ran and time of the walk/run	Track hourly sum of the number of meters walked
Light	Track data about surrounding light levels all the time	Track hourly average of surrounding light level

Noise	Track data about surrounding noise level all the time. Does not record audio	Track hourly average of surrounding noise levels
Weather	Track weather conditions like temperature, pressure, and humidity of the current place all the time	Hourly average of temperature, pressure, and humidity of the environment
Battery	Track battery percentage all the time	Track hourly average battery percentage
Music	Track data about the name of the song for every played song	Track hourly count of the number of songs played
Motion	Track data about device movement, rotation, and magnetic field all the time	Track hourly average of device movement, rotation, and magnetic field
Screentime	Track data about app usage time for every app used on a daily basis	Track data about weekly average screen time
Notification	Track data about the number of notifications received and the name of the app on a daily basis	Track weekly count of the number of notifications received and app names
Pickups	Track data about count on the number of times the phone is picked up from rest, and first opened the app after pickup on a daily basis	Track data about count on the number of times the phone is picked up from rest, and first opened the app after pickup on a daily basis

Table 5.4 Descriptive statistics for feature ratings for tracking Fine-grained data tracking versus Summary data tracking

Fine-grained data tracking					**Summary of data tracking**			
Feature	N	Mean	Std. Deviation	Std. Error Mean	N	Mean	Std. Deviation	Std. Error Mean
Calls	491	3.10	1.242	0.056	491	3.15	1.273	0.057
SMS	491	2.90	1.224	0.055	491	3.00	1.258	0.057
Bluetooth	491	3.07	1.285	0.058	491	3.01	1.313	0.059
Location	491	3.09	1.254	0.057	491	3.12	1.236	0.056
Walk/Run	491	3.80	0.990	0.045	488	3.74	1.010	0.046
Light	491	3.24	1.214	0.055	491	3.22	1.206	0.054
Noise	491	3.18	1.217	0.055	491	3.16	1.235	0.056
Weather	491	3.62	1.076	0.049	491	3.58	1.091	0.049
Music	491	3.11	1.241	0.056	491	3.15	1.272	0.057
Motion	491	3.14	1.211	0.055	491	3.17	1.252	0.056
Screentime	491	3.27	1.170	0.053	491	3.38	1.169	0.053
Notification	491	3.07	1.235	0.056	491	3.12	1.262	0.057
Pickups	491	3.22	1.251	0.057	491	3.21	1.254	0.057

Table 5.5 Paired-Sample test feature ratings for tracking Summary data tracking versus Fine-grained data tracking

Feature	Paired Differences					t	df	Sig. (2-tailed)
	Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference				
				Lower	Upper			
Calls	0.053	1.043	0.047	-0.039	0.145	1.125	490	0.261
SMS	0.102	0.982	0.044	0.015	0.189	2.297	490	0.022
Bluetooth	-0.061	1.096	0.049	-0.158	0.036	-1.236	490	0.217
Location	0.026	0.995	0.045	-0.062	0.115	0.590	490	0.556
Walk/Run	-0.062	0.887	0.040	-0.141	0.017	-1.533	490	0.126
Light	-0.026	0.907	0.041	-0.107	0.054	-0.647	490	0.518
Noise	-0.018	1.009	0.046	-0.108	0.071	-0.403	490	0.687
Weather	-0.041	0.923	0.042	-0.123	0.041	-0.978	490	0.328
Music	0.039	1.065	0.048	-0.056	0.133	0.806	490	0.421
Motion	0.033	0.903	0.041	-0.047	0.113	0.800	490	0.424
Screentime	0.106	1.013	0.046	0.016	0.196	2.317	490	0.021
Notification	0.049	0.959	0.043	-0.036	0.134	1.129	490	0.259
Pickups	-0.012	0.293	0.013	-0.038	0.014	-0.926	490	0.355

There is no significant difference between preference towards one type of tracking i.e., tracking all the time or hourly summary. Overall all the features were rated above the mean value of 2.5. This is a positive sign that people are comfortable to track their mobile sensing data in the context of helping their wellbeing.

To analyze the feature ratings of the app, we conducted a paired-sample t-test and compared the two means for significant differences. The paired-sample t-test shows that there is are significant differences between user's preferences for tracking the features either as fine-grained detail or as the summary for SMS and for Screen time as shown in Table 5.5. Participants prefer summary tracking of SMS (3.0 ± 1.25) as opposed to the Fine-grained data tracking of SMS (2.9 ± 1.22); a statistically significant increase of 0.1018 (95% CI, 0.015 to 0.189), $t(490) = 2.297$, $p < .005$ and prefer summary tracking of Screentime (3.4 ± 1.17) as opposed to the Fine-grained data tracking of Screentime (3.27 ± 1.17); a statistically significant increase of 0.1059 (95% CI, 0.016 to 0.196), $t(490) = 2.317$, $p < .005$. The detailed results of the mean values are as shown in Table 5.4. Again, as shown by the results of the paired-sample t-test on each of the tracking options in Table 5.5, individually, the mean score for all the features of the PROSIT app was higher than the neutral score of 2.5 for the summary tracking option. From the results in Table 5.5, we can see tracking SMS information is the least preferred for both summary and fine-grained tracking. This is in-line with the existing literature about tracking communication being most privacy-invasive [84]. Bluetooth is the second least preferred, this may be because of the potential security vulnerabilities of Bluetooth information [76]. Followed by Notification, Location, and Calls. Not surprisingly, Step count (working or running) is the highest-rated for both tracking methods. This is probably because people are already used to their physical activity being tracked – applications targeted at tracking activity has become so popular. Step count is followed by Weather being second-highest rated. This provides an insight into users' comfortability and privacy perception of mobile sensing features and the importance of tracking summary over fine-grained details of user activities based on user comfortability.

5.1.3 Factors affecting user perception

We measured participants' behavioral *intention* for privacy concerns in 13 scales as listed in Table 5.6.

Table 5.6 Factors affecting privacy perceptions

Abbreviation	**Scale**
PI	Perceived Intrusion
SU	Secondary Use of personal information
PS	Perceived Surveillance
PRC	Privacy Concern
APC	App Permission Concerns
PPB	Privacy Protection Behavior
CA	Computer Anxiety
PPE	Past Privacy Experience
PB	Perceived Benefit
BI	Behavioral Intention
PC	Perceived Control

TR	Trust in app developers and health care providers
INT	Intention to accept app permissions

The questions were measured using a 5-point Likert scale ranging from 1=strongly disagree to 5 strongly agree. We ran Pearson correlation analysis on the 15 sensing features and 13 scales investigating factors affecting the privacy perceptions. Average ratings for All the time and Summary are considered as feature ratings. Table 5.7 below shows the Pearson correlation p-value for the features versus factors. ** represents a significant correlation at 0.01 level.

Table 5.7 Correlation between privacy factors and feature preferences

Factors vs Features	PI	SU	PS	PRC	APC	PPB	CA	PPE	PB	BI	PC	TR	INT
Calls	0.061	0.052	0.049	.217**	.272**	.248**	.325**	.403**	.459**	.497**	.655**	.608**	.588**
SMS	0.058	0.053	.100*	.189**	.253**	.271**	.296**	.370**	.480**	.484**	.639**	.593**	.574**
Bluetooth	0.078	0.043	.099*	.177**	.194**	.221**	.244**	.364**	.443**	.449**	.594**	.544**	.527**
Location	-0.053	-0.003	-0.027	0.086	.156**	.142**	.177**	.308**	.409**	.529**	.567**	.565**	.563**
Steps	0.043	.092*	.120**	0.065	-0.046	0.007	-0.022	0.000	0.040	.260**	0.040	.194**	.181**
Light	-0.026	-0.027	0.003	0.061	.105*	0.078	.152**	.282**	.352**	.413**	.477**	.466**	.504**
Noise	0.016	0.002	-0.025	0.021	0.064	.090*	.123**	.249**	.276**	.384**	.345**	.376**	.434**
Weather	0.000	0.013	-0.026	0.006	-0.056	0.048	-0.024	0.032	.118**	.275**	.152**	.226**	.280**
Battery	0.065	0.035	.102*	.097*	0.054	0.067	.148**	0.055	.120**	.223**	.135**	.171**	.201**
Music	0.061	0.061	0.030	.133**	.160**	.178**	.229**	.243**	.320**	.362**	.438**	.412**	.421**
Motion	-0.033	-0.052	-0.019	0.071	.100*	.095*	.169**	.240**	.375**	.515**	.503**	.521**	.575**
Screentime	0.002	0.015	0.010	.102*	0.081	.151**	.121**	.239**	.333**	.426**	.359**	.391**	.450**
Notifications	0.023	0.018	0.050	.156**	.186**	.204**	.235**	.325**	.395**	.428**	.512**	.514**	.536**
Device	.097*	0.085	.132**	.165**	.137**	.138**	.208**	.224**	.365**	.365**	.394**	.415**	.414**
Pickups	0.042	0.077	0.069	.178**	.222**	.171**	.198**	.317**	.395**	.421**	.434**	.425**	.437**
All-time	0.017	0.033	0.044	.148**	.180**	.200**	.245**	.365**	.457**	.560**	.594**	.605**	.625**
Hourly	0.047	0.046	0.058	.164**	.180**	.197**	.232**	.340**	.453**	.556**	.587**	.592**	.619**
Overall	0.032	0.038	0.050	.158**	.184**	.201**	.241**	.355**	.466**	.568**	.602**	.611**	.635**

PI – Perceived Intrusion, SU–Secondary Use of personal information, PS–Perceived Surveillance, PRC- Privacy Concern, APC–App Permission Concerns, PPB-Privacy Protection Behavior, CA-Computer Anxiety, PPE-Past Privacy Experience, PB-Perceived Benefit, BI-Behavioral Intention, PC-Perceived Control, TR-Trust in app developers and health care providers, INT-Intention to accept app permissions

We observed several significant positive correlations and no significant negative correlation between the factors and features. We looked into non-significant and negative correlations to understand what are the reasons that make people prefer certain features less. Overall Perceived Intrusion (PI) and Secondary use of personal information (SU) did not have a significant positive correlation with any of the tracking features, whereas Perceived Surveillance (PS) had a significant positive correlation with Steps and Device information features while its correlation with all the other features was not significant.

On the other hand, Privacy Concern (PRC) had 10 significant positive correlations and 8 non-significant correlations. App Permission Concerns (APC) had a significant positive correlation with 11 features and a nonsignificant correlation with 7 features. Privacy Protection Behaviour (PPB) was positively correlated with 12 features and had a non-significant correlation with 6 features. Computer Anxiety (CA) had a significant positive correlation with 16 features and had a non-significant correlation with 2 features. Past Privacy Experience (PPE) had a significant positive correlation with 15 features and was not correlated with 3 features. Perceived Benefit (PB) and Perceived Control (PC) was positively correlated with all the features except Steps. Other factors Behavioural Intention (BI), Trust in app developers and health care providers (TR), and Intention to accept app permissions (INT) had a significant positive correlation with all the 18 features. This implies that Perceived intrusion (PI), Secondary use of personal information (SU), and Perceived Surveillance (PS) do not impact (or share any relation with) users' perception of mobile sensing features for tracking for health and wellbeing privacy. On the other hand, Behavioral Intention (BI) to use a mobile sensing app, Trust in app developers and health care providers (TR), and Intention to accept app permissions (INT) are significantly related to user's comfortability of mobile sensing features for tracking health and wellbeing privacy. Hence, they impact the user's perception of mobile tracking privacy.

On average, we asked participants to rate the app's usefulness during a pandemic situation. Participants with a history of mental health illness provided an average rating of 3.8 and participants without a history of mental health illness provided an average rating of 3.6. This shows that people who had mental health illness in the past perceive the PROSIT app to be more useful during a pandemic. However, healthy controls also perceived the app as useful as they rated it above the mean rating of 2.5.

5.2 PROSIT app pilot study

The app was used by the PROSIT research team, the participant recruitment started in November 2019. Over 200 participants were recruited to use the iOS application. The app performed well in different apple phone models, OS versions and continuously collected and uploaded 40 million records as of May-2020. Technical issues concerning local DB locks caused the app to crash initially, and the numbers were brought down in the subsequent versions. The app had 6 major releases and was distributed via TestFlight [133]. To evaluate the quality and usefulness of data collected, a dataset was obtained from the PROSIT research team. The dataset did not contain any identifying information about the participants, each participant is identified by a generated PROSIT ID. Initial analysis was conducted on the obtained dataset. The analysis steps are Pre-processing, exploratory analysis, principal component analysis, and classification. The results of this analysis are presented in this section.

5.2.1 Dataset information

Out of 115 participants in both iOS and Android, a dataset of 90 iOS participants and labels were obtained from the PROSIT research team for this preliminary analysis. This is because iOS is the understudied platform for Mobile sensing applications for mental health. Participants with a minimum of 10 days of data before 15-March-2020 were shortlisted to avoid bias due to the COVID-19 lockdown imposed on 15-March-2020 in Nova Scotia. Based on the labeling information 0 or 1 for each participant the dataset is labeled in two groups healthy subjects and patients respectively. Only participants with GPS location coordinates in and around Nova Scotia are considered. The location filtering is done to eliminate any bias due to COVID-19, as lockdown dates and impact may differ in different parts of the world. After filtering based on all the above-mentioned criteria, 18 participants were shortlisted, with equal distribution of 9 participants from healthy subjects and 9 participants from patients. The number of participants in both groups is set to be equal to eliminate bias.

5.2.2 Pre-processing

The linear data was converted into dataset format by aggregating the features. Event-based features were aggregated by the count method and numeric features are aggregated by calculating mean. Data is aggregated in hourly intervals. The final Dataset consisted of 27 columns, with no missing information, and contained an equal number of participants for targets 0 and 1 indicating healthy subjects and patients respectively. The final dataset for analysis had the groups identified by numbers 0 or 1, an hour from 0 to 23 followed by feature values. All other information like participant IDs, dates, and timestamps were removed for privacy.

The feature abbreviations used in column labels and corresponding description is listed in the table below

Table 5.8 Feature description

Index	Feature label	Description
0	hour	Hour of the day ranges from 0 to 23
1	accelX	Accelerometer X value
2	accelY	Accelerometer Y value
3	accelZ	Accelerometer Z value
4	brightness	Mobile phones' screen brightness
5	callIncoming	Incoming voice/video call
6	callOutgoing	Outgoing voice/video call
7	callDisconnected	Call disconnected
8	callConnected	Call connected or answered
9	gyroX	Gyroscope X value
10	gyroY	Gyroscope Y value
11	gyroZ	Gyroscope Z value
12	latitude	Latitude of current GPS location
13	longitude	Longitude of current GPS location
14	locked	Screen locked
15	unlocked	Screen unlocked
16	magX	Magnetometer X value
17	magY	Magnetometer Y value
18	magZ	Magnetometer Z value
19	batteryPercent	Battery charge percentage
20	rchWifi	WiFi connectivity

21	rchCellular	Cellular or Mobile data connectivity
22	rchNo	No internet connectivity
23	sleepNoise	Noise levels recorded during sleep
24	temperature	Atmospheric Temperature of current location
25	pressure	Atmospheric Pressure of current location
26	humdity	Humdity percentage of current location
27	group	Target group 0 or 1

5.2.3 Exploratory Data Analysis

Exploratory data analysis was performed by plotting pairwise box plots and scatter plots for each of the 26 features listed in Table 5.14. Correlation coefficients were calculated for all the 26 features and plotted in a heat map. Some of the significant positive correlations were observed within the motion sensors magnetometer, gyroscope and accelerometer, screen lock and unlock states has a positive correlation with motion sensors and call events but negatively correlated with location changes. Wifi connectivity state is negatively correlated with motion sensors, this can be related to the general fact that WiFi is available within buildings where location changes are less likely to happen. In contrast, rchNo and rchCellular state is positively correlated with motion sensors, this can be related to the unavailability of Wi-Fi outdoors, which requires mobile users to connect to mobile data or be offline. Weather features temperature, pressure, and humidity is negatively correlated with rchNo state, this is in line with the fact that weather information retrieved from the internet cannot be obtained when the user is offline.

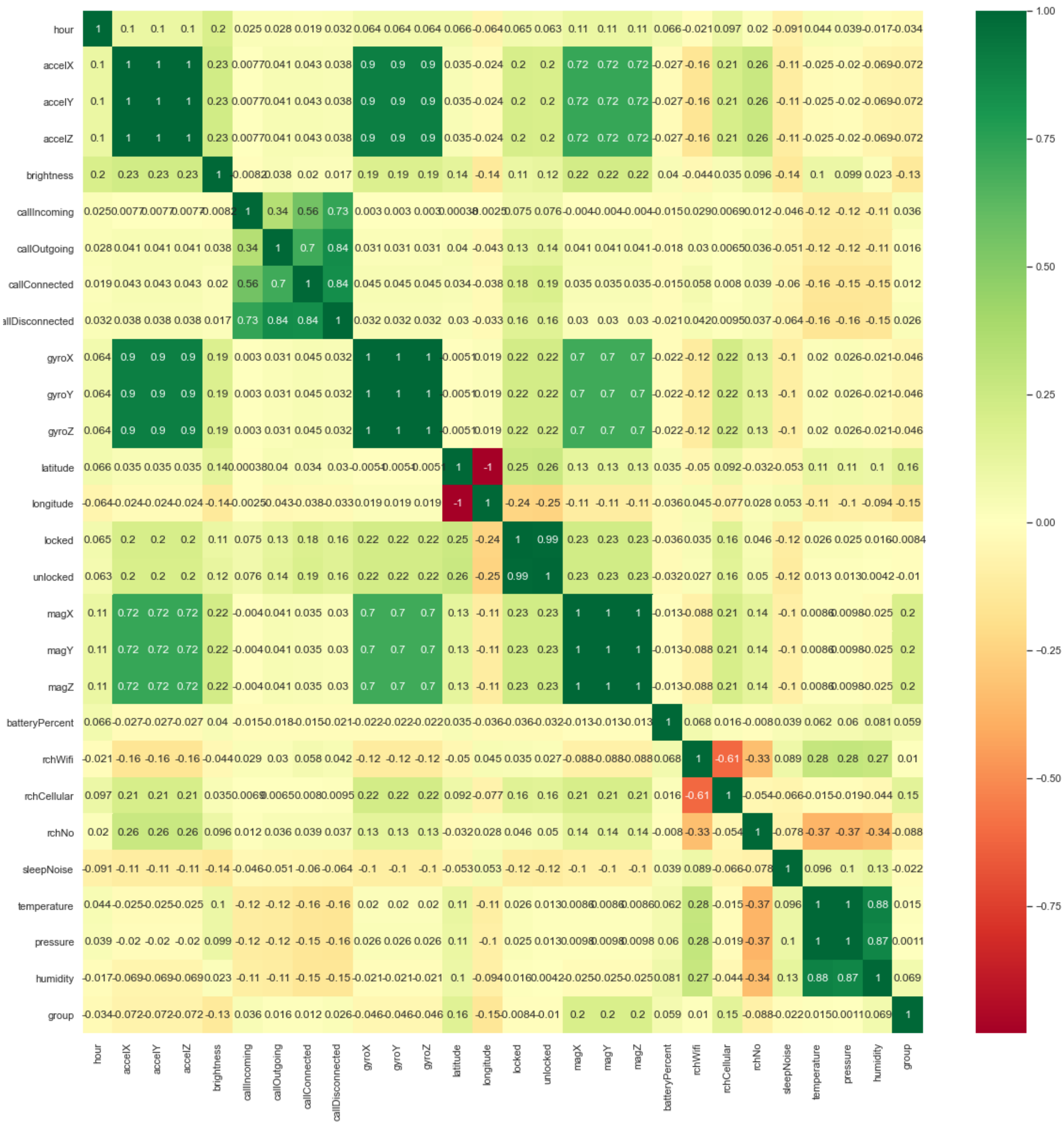


Figure 5.1 Correlation heatmap

The Correlation heatmap, Figure 5.1, shows significant positive correlations between the variables, which is in line with the ground truth information taken into account while developing and testing the PROSIT application.

5.2.4 Principal Component Analysis

To understand the features that distinguish the target classes healthy and sick, we calculated feature importance using the sklearn PCA library. The results show

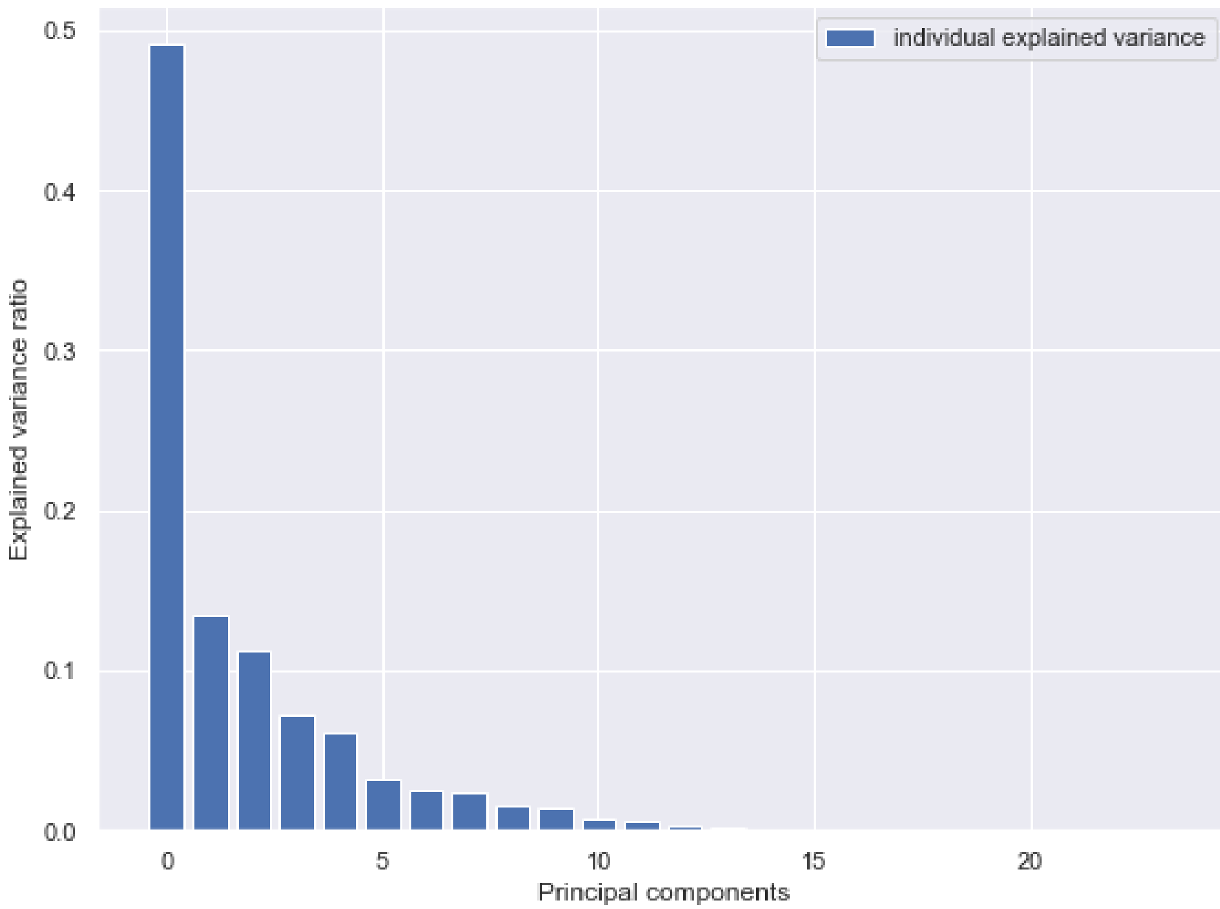


Figure 5.2 PCA Explained variance ratio

The highest variance of Principal Component 1 (PC1) is 49%, followed by Principal Component 2 (PC2) at 13% and Principal Component 3 (PC3) at 11%. The first three principal components added up to provide 73% of variance explained by the features. The 3-Dimensional scatter plot below shows the distribution of the groups along three axes PC1, PC2, and PC3.

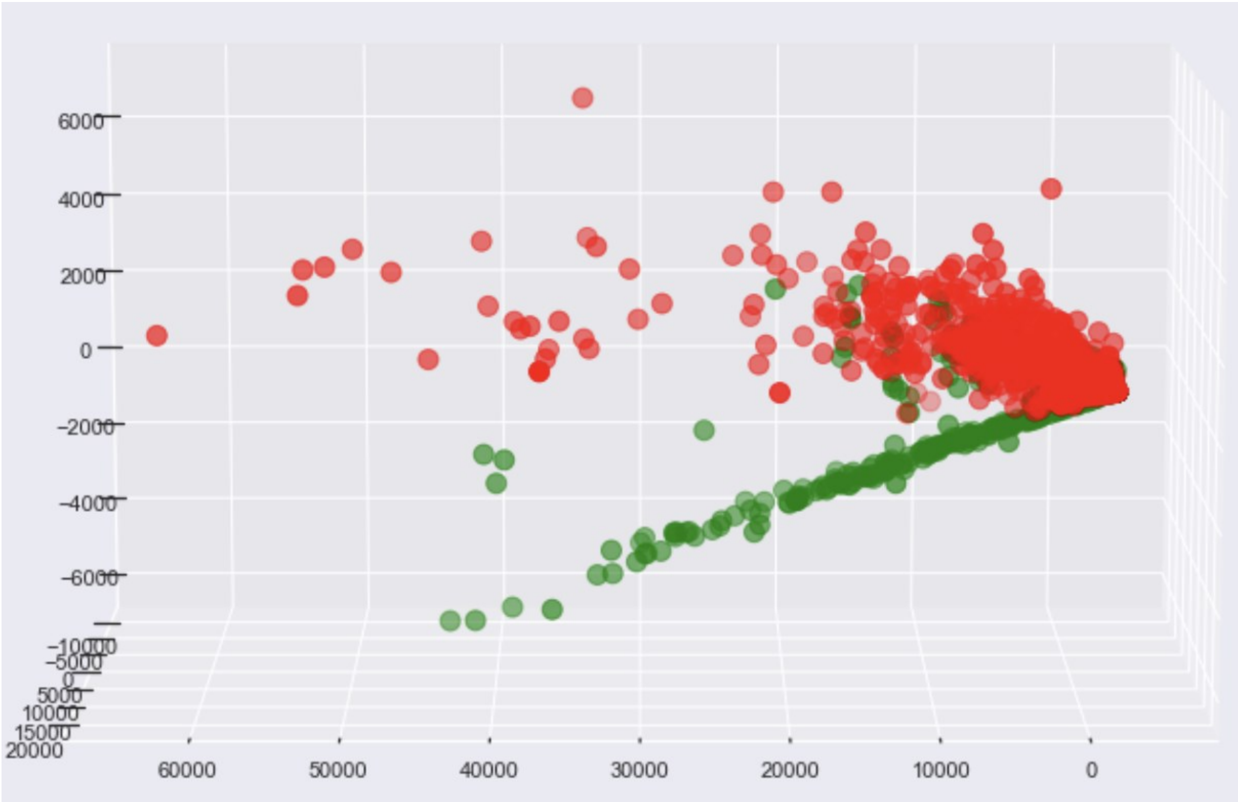


Figure 5.3 K-nearest neighbor classifier result 3D scatter plot

To understand the features contributing to the entropy, the ExtraTreeClassifier algorithm is executed to rank the important features. From the results, we could see the motion sensors and screen brightness contributed to 50% of the ranks.

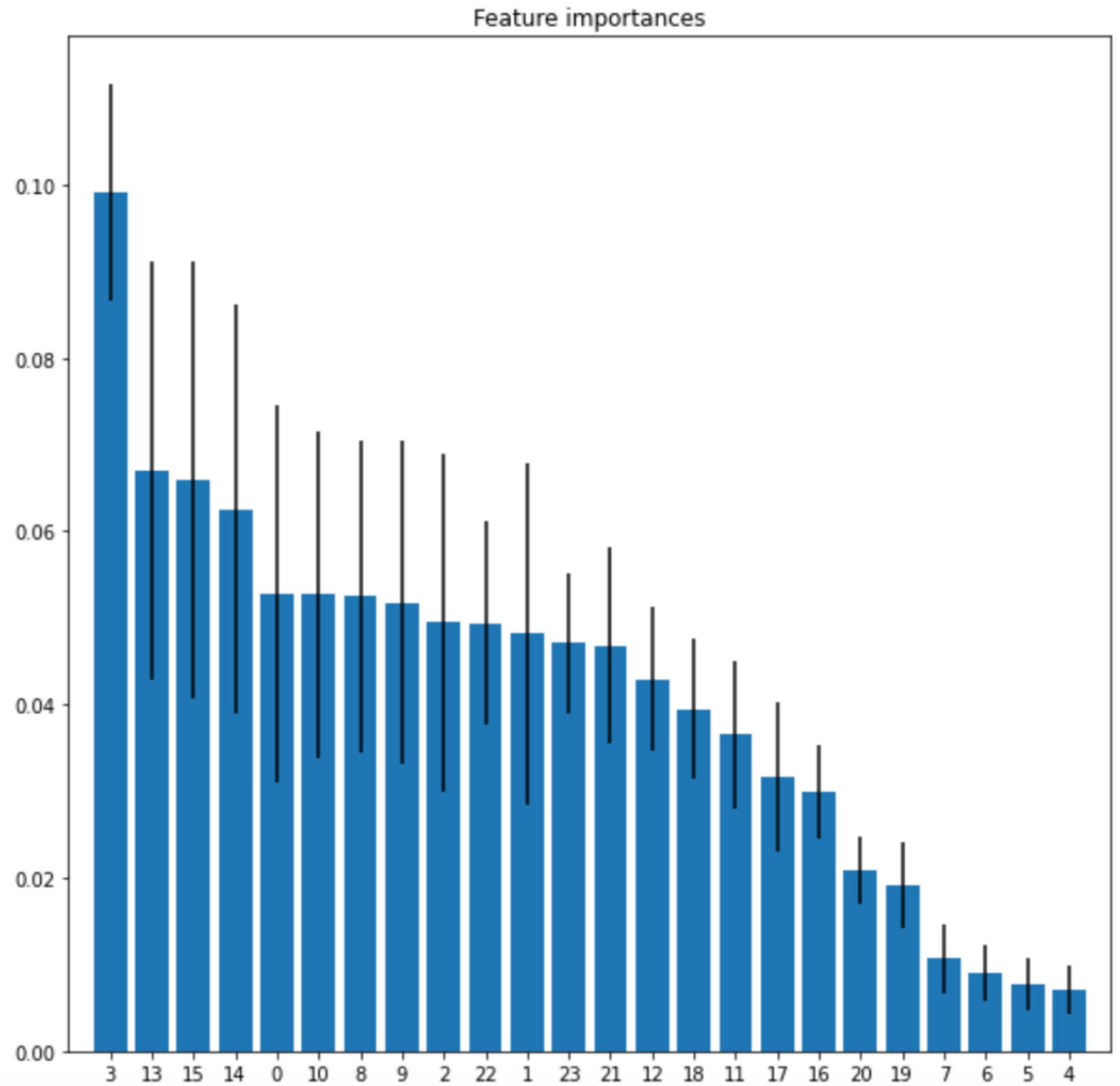


Figure 5.4 Extra Trees Classifier feature importance

Table 5.9 Feature ranking and importance scores

Rank	Index	Importance	Feature
1	3	0.099099	brightness
2	13	0.066956	magX
3	15	0.065915	magZ
4	14	0.06251	magY
5	0	0.052816	accelX
6	10	0.05273	gyroZ
7	8	0.052493	gyroX
8	9	0.051781	gyroY
9	2	0.04948	accelZ

10	22	0.049388	pressure
11	1	0.048125	accelY
12	23	0.047066	humidity
13	21	0.046804	temperature
14	12	0.042936	unlocked
15	18	0.039478	rchCellular
16	11	0.036518	locked
17	17	0.031665	rchWifi
18	16	0.029911	batteryPercent
19	20	0.020869	sleepNoise
20	19	0.019044	rchNo
21	7	0.010693	callDisconnected
22	6	0.008987	callConnected
23	5	0.007723	callOutgoing
24	4	0.007013	callIncoming

In smartphones, brightness is directly proportional to ambient light. Participants of Group healthy participants have higher brightness levels as compared to Group of patients.

5.2.5 Anomaly detection algorithm

The anomaly detection algorithm feasibility study results indicate significant improvement in terms of memory usage, power consumption, and battery usage among the versions PROSIT and PROSITLite.

5.2.6 Power consumption and Memory usage

The power consumption results show that the anomaly detection method uses significantly low power as compared to raw data collection. From the pilot study, we observed trends in Wi-Fi connectivity, and the battery charge is high during night-time. This indicates people tend to charge their devices at night and connected to Wi-Fi. As PROSIT is aware of Wi-Fi connectivity and charging state, we can utilize this time to run the algorithm. On average PROSIT consumes 24% of the device battery on iPhones whereas PROSITLite consumed less than 8% of the battery. This is equivalent to a 67% reduction in battery consumption by PROSITLite.

From the results of two weeklong study on PROSIT and PROSITLite, we calculated averages in terms of memory and power consumption. The results are as presented in Table 5.10. The

PROSITLite lead to a 97% reduction in storage space (memory usage) and a 67% reduction in power consumption.

Table 5.10 Comparison of PROSIT and PROSITLite

Factors	PROSIT	PROSITLite
Average memory usage (KB)	3847	125
Average power consumption (%)	24	8

In addition to the reduction in the amount of memory and power consumption, Federated Learning also saves significant time on data analysis and prediction, thus helping clinicians to make informed decisions in a shorter period. Most importantly, it preserved user's privacy since most data do not have to leave users' phones to the server as it would in the PROSIT.

5.3 Discussions of the Results

In this research. We first reviewed literature and identified gaps to be lack of mobile sensing applications for the iOS operating system, collaboration with psychiatry, and on average, available apps sense between 1 to 10 different type of sensors. To fill these gaps, we designed, implemented, and conducted a preliminary evaluation of the PROSIT app, as a mobile application for both iOS and Android with all possible tracking features listed in the literature, few features were removed after discussing with the psychiatrist.

We evaluated the PROSIT app to understand three different aspects of Mobile sensing applications for the Mental Health Monitoring System (MHMS). Firstly, we conducted an online survey to understand the privacy concerns related to data tracking using mobile sensing and the factors that affect the privacy concerns of the users. Secondly, we conducted a pilot study of 18 participants (9 patients and 9 health subjects) who used the app for two weeks. We analyzed the results of the pilot study for validating the app for usability, and data quality to classifying patients and non-patients. Finally, we implemented the Federated Learning algorithm for the PROSIT app to improve the overall performance of battery consumption, energy consumption, data storage space (memory usage), privacy-preserving, and the usefulness of the summarized data. We conducted a feasibility study to compare and understand the efficiency and benefits.

5.3.1 Privacy score of Mobile Sensing features

The mean values for all the variables measured under feature ratings were 2.5 as seen in Table 5.17, therefore we accept the alternative hypothesis which states that the users of mobile sensing applications for health and wellness have less or no privacy concerns in terms of tracking their mobile sensing data for their wellbeing. Of the 15 feature ratings measured, all of these features showed a positive rating because they were greater than the neutral value of 2.5 as shown in Figure 5.7. However, in general, people preferred summary tracking compared to fine-grained tracking.

Table 5.11 Mean rating and Standard deviation for features

Feature	Mean	Standard Deviation
SMS	2.94	1.13
Bluetooth	3.03	1.17
Notifications	3.08	1.15
Location	3.10	1.14
Calls	3.12	1.14
Music	3.13	1.14
Motion	3.14	1.14
Noise	3.16	1.11
Pickups	3.20	1.24
Light	3.22	1.12
Screentime	3.32	1.05
Battery	3.44	1.06
Weather	3.59	0.98
Steps	3.76	0.89

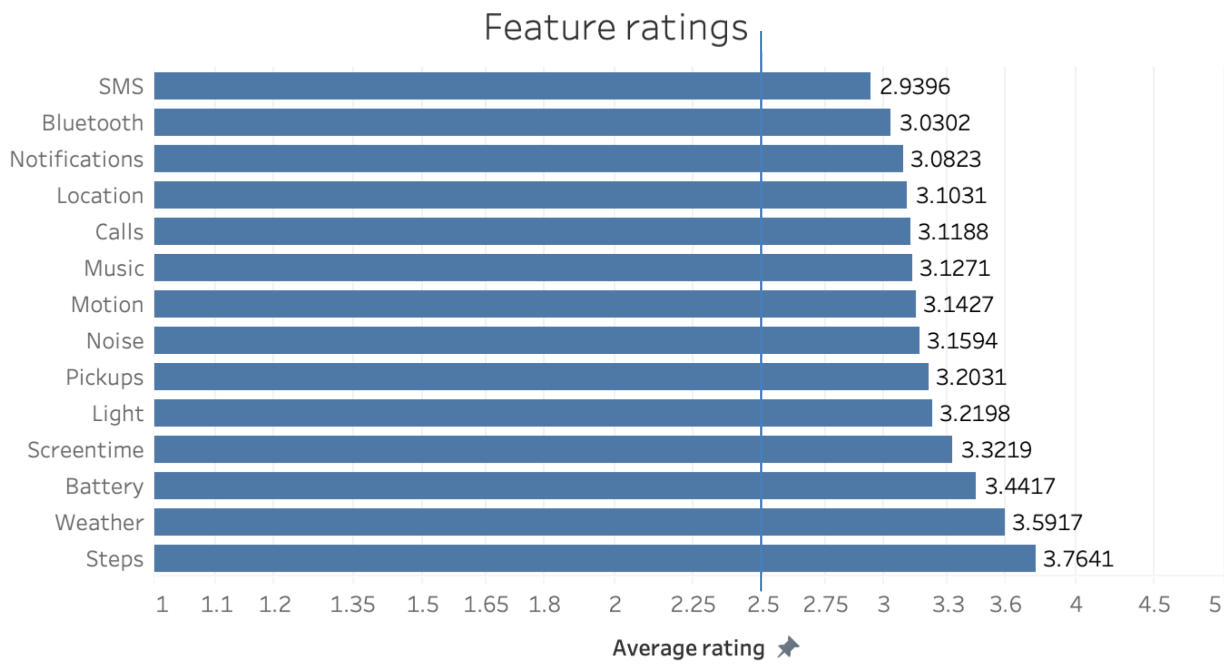


Figure 5.5 Average ratings of tracking features

Since all the feature ratings are positive, we can conclude that the users have a positive attitude towards tracking mobile sensing data for health and wellbeing in general. However, users are more critical of some features than the others. Features such as SMS, calls, Bluetooth, Location, and Notifications were not rated significantly high above neutral. Hence, users not as unconcerned about their privacy when this featured being tracked in detail. This is understandable considering the level of information that can be revealed by these data if not properly managed. For summary tracking, users not as unconcerned about their privacy when SMS and Bluetooth information are being tracked. This shows that it is generally preferred to track only summarized user data than the fine-grained detailed data of users.

5.3.2 Factors affecting privacy perception

The correlation analysis for all the thirteen factors measured against user's privacy comfortability rating for all the features was non-negative, therefore we can conclude that the factors related to privacy perceptions either positively impact the feature ratings related to mobile sensing or has no impact at all. Table 5.12 provides the correlation values of the 13 factors, with ** representing significant correlation at the 0.01 level.

Table 5.12 Factors affecting mobile sensing feature ratings - Correlation

Factors	**Description**	**Pearson Correlation**	**Sig. (2-tailed)**
PI	Perceived Intrusion	0.032	0.4845
SU	Secondary Use of personal information	0.038	0.3993
PS	Perceived Surveillance	0.050	0.2735
PRC	Privacy Concern	.158**	0.0004
APC	App Permission Concerns	.184**	0.000
PPB	Privacy Protection Behaviour	.201**	0.000
CA	Computer Anxiety	.241**	0.000
PPE	Past Privacy Experience	.355**	0.000
PB	Perceived Benefit	.466**	0.000
BI	Behavior Intention	.568**	0.000
PC	Privacy Concern	.602**	0.000
TR	Trust in app developers and health care providers	.611**	0.000
INT	Intention to accept app permissions	.635**	0.000

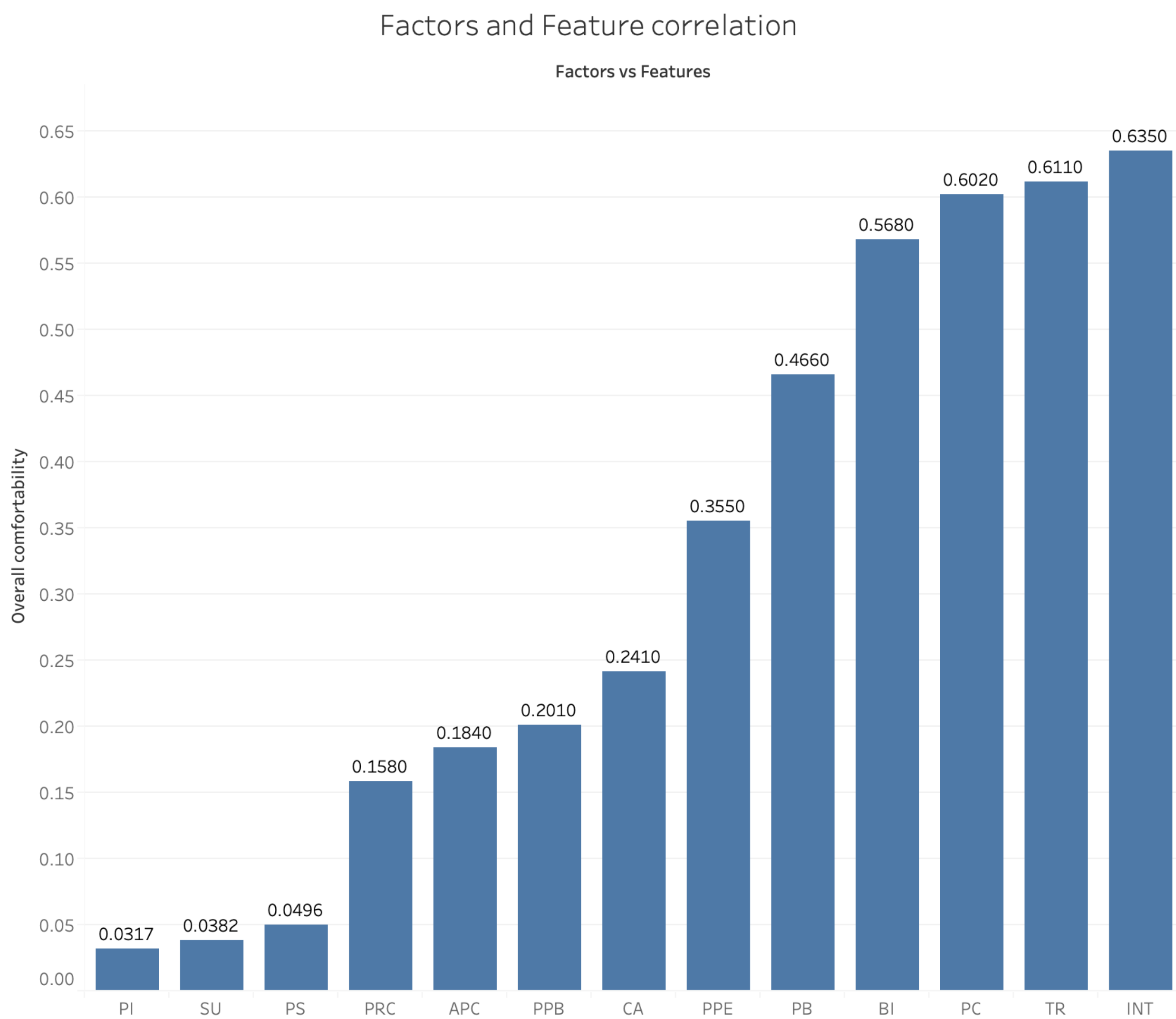


Figure 5.6 Factors affecting mobile sensing feature privacy comfortability ratings

From Figure 5.6 we could see the statistical significance of each of the thirteen factors measured versus the average rating of all the mobile sensing features. The correlation coefficient of Perceived Intrusion (PI) is the lowest with r value 0.0317, followed by Secondary Use of Personal information (SU) r value 0.0382 and Perceived Surveillance (PS) with r value 0.0496. The above three factors do not have a significant positive correlation with the user's intention. This concludes that users concern towards mobile sensing apps are about Intrusion in terms of their private information being made publicly available without consent, Secondary Use by which the users' data is used for other purposes that they did not provide consent, and Surveillance where the user perceives that a mobile sensing app collects too much information by monitoring everyday activities. All the other ten factors have a significant positive correlation with the mobile sensing feature tracking.

5.3.3 PROSIT as Mental Health Monitoring System

To recap the results, of our pilot study. The dataset with 27 features was analyzed to find the correlation among the features. On a high level, these 27 features are grouped into 12 features namely *Brightness, Call event, Location tracking, Lock event, Motion, Battery percentage, Wi-Fi connectivity, Mobile data connectivity, Sleep tracking,* and *Weather*. The information about each feature is as follows. *Brightness* values represent values from 0 to 0.1 representing smartphone screen brightness value, which turns 0 when the screen backlight is turned off and has a non-negative value when a screen interaction is taking place. *Call event* has combined count values four types of call events namely Incoming call, Outgoing call, Call connected, Call disconnected. *Location tracking* has a derived binary value 0 or 1, to represent whether location tracking is turned ON or OFF. *Lock event* is derived from combining the number of smartphone locks and unlocks. *Motion* feature is derived from 9 other features from Accelerometer, Gyroscope and Magnetometer values in x, y and z axes, the magnitude of each sensor is calculated and finally added to derive the *Motion* feature. *The battery percentage* is a feature the represents the average smartphone battery level. *Wi-Fi connectivity* represents the smartphone connected to the Wi-Fi, Mobile data connectivity represent the smartphone connected to mobile data, which implies the device uses the internet provided by the SIM card service provider. *Sleep tracking* represents the manual tracking feature of the PROSIT app, the users are requested to turn on Sleep noise tracking before going to bed and turn it off after the person wakes up in the morning, *Weather* feature is derived from Temperature, Pressure, and Humidity sensing values of the dataset, it indicates whether the tracking is enabled or not, Weather tracking relies on location services i.e current latitude and longitude coordinates of the smartphone to obtain location-specific weather information of the user. *Hour* feature is derived from the hour of the day, it holds one of 3 values from the set {0,1,2} where each value represents different parts of a day. Value 0 represent *Night* from 12:00 AM to 09:00 AM, Value 1 represent *Evening* from 06:00 PM to 11:59 AM, Value 2 represent *Daytime from* 10:00 AM to 06:00 PM. The results of the correlation are aligned with the ground truth indices about phone usage and routines and are described in Table 5.19

Table 5.13 Correlation analysis between features and ground truth

Feature	Correlated with	Type of significant correlation (0.01 level)	Inference with ground truth
Brightness	Hour	Positive	Brightness increases during daytime
	Call event	Positive	Brightness increases when there is a call event. Connecting and Disconnecting a call are related to touch screen interaction.
	Sleep tracking	Negative	The phone is not used while the user is sleeping. Since there is no screen interaction, the brightness goes down when the user manually turns on Sleep tracking and goes to bed.
Call Event	Hour	Positive	More calls are made during the daytime.
	Motion	Positive	There are movements related to hand motion or walking when the user is has a call event.
	Lock event	Positive	The mobile phone is locked or unlocked during call events. The events are unlocking to make a call, Locking the phone after disconnecting a call.
	Sleep tracking	Negative	Less or no calls are made when the user is sleeping
Lock event	Sleep tracking	Negative	A smartphone screen is not frequently locked/unlocked while the user is sleeping
Battery percentage	Hour	Negative	Smartphone battery is drained during the daytime when the usage is high
	Sleep tracking	Positive	Users tend to charge their phones at night when the usage is less
Wi-Fi connectivity	Hour	Negative	The smartphone is connected to Wi-Fi (Home Wi-Fi) during night-time
	Sleep tracking	Positive	
Mobile data connectivity	Hour	Positive	The smartphone is connected to Mobile data during the daytime, as the user is not connected to Home Wi-Fi.
	Sleep tracking	Negative	The smartphone is not connected to mobile data during bedtime, since users are at home connected to Home Wi-Fi
Location tracking	Hour	Positive	Location changes occur during daytime

Weather	Location tracking	Positive	Weather information is acquired based on user location, the data is acquired only when location services are turned on.

Based on the results from the correlation analysis, we can conclude that the PROSIT app is efficient in tracking daily indices of the app user, by passively tracking mobile sensing information in the background.

5.3.4 Federated Learning framework for detecting Depression

From the results of the simulation study, it is evident that Federated learning algorithm implementation has significantly reduced the client-side overhead in terms of battery power consumption and memory utilization. From Figure 5.7. On average PROSIT consumes 24% of phone battery whereas PROSITLite consumes 8% of battery.

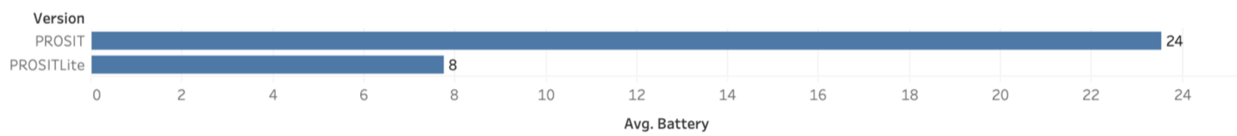


Figure 5.7 Battery consumption of PROSIT and PROSITLite

From Figure 5.8 we can see a significant difference in memory utilization between PROSIT and PROSITLite. On average PROSIT takes up 3137 KB of storage space whereas PROSITLite takes 45 KB of memory. This significant reduction in storage is a result of discarding the mobile sensing data after the median values are computed for the anomaly detection algorithm.

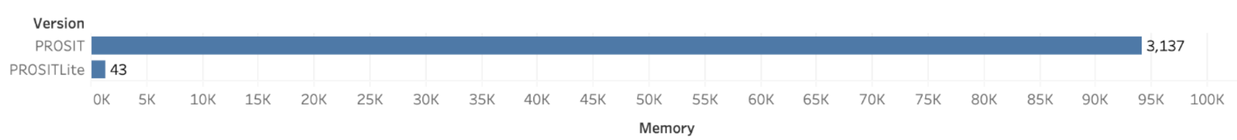


Figure 5.8 Memory consumption PROSIT and PROSITLite

The reduction in memory consumption is also due to PROSIT tracking 23 types of mobile sensing data to detect various mental health disorders, while PROSITLite is restricted to tracking only three types of mobile sensing information such as location, motion, and calls which are key indicators for detecting depression. The reduction in battery and memory usage concludes that PROSITLite is efficient in terms of optimal utilization of client-side resources.

Since the COVID-19 pandemic and related measures have brought significant changes to the day-to-day life of participants, such as limited outdoor movement resulting in fewer location changes, we could not test the accuracy of the anomaly detection algorithm to detect the onset of depression.

5.3.5 Clinical Blurb from Psychiatrist

Currently, 1 in 5 youth around the world is affected by a mental disorder [45]. Although there are many evidence-based treatments available, less than 20% of youth suffering from mental disorders receive appropriate care [45]. Also, 20%-40% of youth do not respond sufficiently to the treatments they receive resulting in continued suffering and chronicity [142][67]. If health professionals can't predict and efficiently prevent new onsets and poor outcomes of mental disorders in youth, the field faces a crisis. Particularly, intervening at critical moments - that is, during mental health crises, including times of risk for suicide, self-harm, psychotic breakdown, substance use relapse, and interpersonal loss can have a major impact on improving youth's mental health. Current methods of predicting mental health crises usually rely on subjective symptom ratings obtained at discrete time points during routine clinical care (e.g., clinical monitoring, screening). But clinical decision-making based on such subjective information is challenging, as changes in symptoms might be sub-threshold, context-dependent, or too subtle to be captured using subjective patients' ratings. The PROSIT tool is designed to capture multiple indices of youth's daily life behaviors in an unobtrusive manner. The indices are aligned with findings that demonstrate their potential to make inferences about their mental health status [113][49]. These indices include physical activity, geolocation, sleep, phone use, typed text, music choice, and acoustic vocal quality. I do believe the tool can, if further validated, provide health care professionals and youth patients with actionable insights to reduce distress and prevent onset of mental disorders and poor treatment outcomes. In addition, in my experience youth patients are further more compliant with passive clinical monitoring (i.e., via the PROSIT tool) than active weekly or daily symptoms rating.

5.3.6 Future work in Anomaly detection and Federated Learning

In the future, we plan to run a long-term study on depressive patients to validate the accuracy of the anomaly detection algorithm in detecting the onset of a depressive episode.

Federated Learning is in its infancy, software libraries have limited support, and most research in this area in a conceptual stage. CORE ML is the Apple Inc provided library [31] that is used to implement on-device machine learning features in iOS applications. The latest version 3 supports on-device training of Image and Audio data, and Tabular data is not supported yet. Since. The Anomaly detection algorithm proposed in this thesis can be used in the future to label and train datasets on the mobile device, we plan to complete the client-server implementation of Federated Learning architecture and evaluate the efficiency.

CHAPTER 6 CONCLUSION

In this chapter, we summarized the thesis and highlight the limitations, contributions, and suggested potential directions for future work.

6.1 Limitations

The main limitation of the research is the subjective nature of the survey responses. Although in the survey we provided clear descriptions about each sensing feature, and its purpose and instructed the participants to answer the questions sincerely, it is common knowledge that human perception is not always perfect, and bias would most likely be present. There is no way to ensure that participants were understanding the mobile sensing feature concepts before answering the questions although we included attention questions to ensure participants were reading and considering their responses carefully. Nevertheless, self-report is still a valid and predominant approach for assessing users' opinions.

The pilot study is on a small group of 18 participants, hence the results and classification accuracy is limited to the data from the 18 participants. The 9 patients in the pilot study were classified based on lifetime diagnosis. The prediction will get better with actual dimensional symptomatology. Since there is limited research focusing on iOS users, the 18 participants were chosen to be all iOS users. Further analysis is required on Android users from the dataset collected from the Android version of the PROSIT app. The limitation of a small dataset will be eliminated when the ongoing study is complete.

Federated Learning is a new framework, and hence software libraries are limited. Apple provided COREML 3 only supports training Image and Audio data on the device, whereas tabular data training is not supported. This limits the complete implementation of the Federated learning framework. If on-device training for tabular data is offered by Apple, we can use the anomaly detection algorithm to label and retrain models on the device.

6.2 Contributions

We successfully designed and developed a mobile sensing application that can, continuously collect mobile sensing information and periodically transfer it to secure servers. The privacy, security, and battery consumption aspects are optimized during the design and development phases, and the app is built in close collaboration with the Psychiatry domain expert. This closes the research gap that has been existing in this area. We successfully conducted a pilot study and the results are promising, showing that PROSIT will be helpful to run a large-scale study and fit for future inclusion in clinical practice. Finally, the simulated results of the Federated learning approach show a promising approach for long-term studies, further eliminating drawbacks in current mobile sensing applications for mental health.

This research has gained significant media attention. It has been featured on CBC news [33], Body and Soul [109], and 40 other media outlets in 16 different countries.

6.3 Future work

In the future, we plan to thoroughly train a model with data from an ongoing study with more than 300 participants, and other publicly available datasets like *StudentLife* [130]. This comprehensive model is planned to be implemented in a Federated Learning framework, for on-device learning and efficient patient monitoring.

Additionally, we have already built and tested features like Music tracking, a custom keyboard, integration with wearables to capture health indices and push notifications for intervention. We also built a user interface to capture the daily mood of the participants. These additional features will be used in future studies for in-depth analysis of mobile sensing for mental health.

6.4 Conclusion

This thesis contributes to an important research field of Mobile sensing within the Mental health (mHealth) domain by developing a mobile sensing application for mental health that supports iOS and Android platforms. It unobtrusively tracks more than 15 types of sensing information from smartphones thus capturing daily indices of people and periodically transfers the data to clinical servers. The results from the survey show that our participants showing positive attitudes towards mobile sensing, especially during pandemic times like COVID-19 when mental health care services such as in-person clinical visits are not available for the greater population. Mobile users like to use sensing apps if they trust the application developers and markers if the users are given control to accept application permissions and have control over the data being shared. On the other hand, users do skeptical about mobile sensing for health due to possible surveillance, intrusion, and the data getting used for other purposes. Users with personality type Neuroticism and Openness like PROSIT app and personality type Conscientiousness doesn't prefer the app. The result from the pilot study on 18 participants validated the quality of data collected, and its ability to classify two groups of participants with 73% accuracy using a clustering algorithm. More importantly, we implemented and tested an anomaly detection algorithm that lay the foundation for Federated Learning for PROSIT, the approach was effective in reducing several overheads in terms of memory, power consumption, and transferring large datasets to the server, and then performing data analysis for months. Hence, this app can be used to definitively monitor and provide interventions for patients with mental health disorders.

BIBLIOGRAPHY

[1] • Mobile OS market share 2019 | Statista: *https://www.statista.com/statistics/272698/global-market-share-held-by-mobile-operating-systems-since-2009/*. Accessed: 2020-08-01.

[2] • Mobile OS share in North America 2018-2020 | Statista: *https://www.statista.com/statistics/1045192/share-of-mobile-operating-systems-in-north-america-by-month/*. Accessed: 2020-11-25.

[3] • Smartphone users worldwide 2020 | Statista: *https://www.statista.com/statistics/330695/number-of-smartphone-users-worldwide/*. Accessed: 2020-08-05.

[4] 325,000 mobile health apps available in 2017 – Android now the leading mHealth platform: 2017. .

[5] A Study on the Development of a Day-to-Day Mental Stress Monitoring System using Personal Physiological Data - IEEE Conference Publication: 2018. *https://ieeexplore.ieee.org/document/8571973*. Accessed: 2020-08-02.

[6] Andrews, S. et al. 2015. Beyond self-report: Tools to compare estimated and real-world smartphone use. *PLoS ONE*. 10, 10 (Oct. 2015). DOI:https://doi.org/10.1371/journal.pone.0139004.

[7] Android v iOS market share 2019: *https://deviceatlas.com/blog/android-v-ios-market-share#sweden*. Accessed: 2020-08-01.

[8] Balsari, S. et al. 2018. Reimagining health data exchange: An application programming interface-enabled roadmap for India. *Journal of Medical Internet Research*. 20, 7 (Jul. 2018), e10725. DOI:https://doi.org/10.2196/10725.

[9] Baras, K. et al. 2016. “Smartphine”: Supporting students’ well-being according to their calendar and mood. *2016 International Multidisciplinary Conference on Computer and Energy Science, SpliTech 2016* (Aug. 2016).

[10] Barnett, I. et al. 2018. Relapse prediction in schizophrenia through digital phenotyping: A pilot study. *Neuropsychopharmacology*. (2018). DOI:https://doi.org/10.1038/s41386-018-0030-z.

[11] Ben-Zeev, D. et al. 2015. Next-Generation psychiatric assessment: using smartphone sensors to monitor behavior and mental health. *Psychiatric Rehabilitation Journal*. 38, 3 (Sep. 2015), 218–226. DOI:https://doi.org/10.1037/prj0000130.

[12] Berrouiguet, S. et al. 2018. Combining Continuous Smartphone Native Sensors Data Capture and Unsupervised Data Mining Techniques for Behavioral Changes Detection: A Case Series of the Evidence-Based Behavior (eB2) Study. *JMIR mHealth and uHealth*. 6, 12 (Dec. 2018), e197. DOI:https://doi.org/10.2196/mhealth.9472.

[13] Berto, P. et al. 2000. Depression: cost-of-illness studies in the international literature, a review. *The Journal of Mental Health Policy and Economics*. 3, 1 (Mar. 2000), 3–10. DOI:https://doi.org/10.1002/1099-176X(200003)3:1<3::AID-MHP68>3.0.CO;2-H.

[14] Bonawitz, K. et al. 2019. Towards Federated Learning at Scale: System Design. (Feb. 2019).

[15] Boonstra, T.W. et al. 2017. Smartphone app to investigate the relationship between social connectivity and mental health. *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS* (Sep. 2017), 287–290.

[16] Boonstra, T.W. et al. 2018. Using Mobile Phone Sensor Technology for Mental Health Research: Integrated Analysis to Identify Hidden Challenges and Potential Solutions. *Journal of medical Internet research*. 20, 7 (Jul. 2018), e10131. DOI:https://doi.org/10.2196/10131.

[17] Boukhechba, M. et al. 2017. Monitoring social anxiety from mobility and communication patterns. *UbiComp/ISWC 2017 - Adjunct Proceedings of the 2017 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2017 ACM International Symposium on Wearable Computers* (New York, NY, USA, Sep. 2017), 749–753.

[18] Boukhechba, M. et al. 2017. Monitoring social anxiety from mobility and communication patterns. *UbiComp/ISWC 2017 - Adjunct Proceedings of the 2017 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2017 ACM International Symposium on Wearable Computers* (New York, NY, USA, Sep. 2017), 749–753.

[19] Brendan McMahan Eider Moore Daniel Ramage Seth Hampson Blaise AgüeraAg, H. and Arcas, A. 2017. *Communication-Efficient Learning of Deep Networks from Decentralized Data*.

[20] Brisimi, T.S. et al. 2018. Federated learning of predictive models from federated Electronic Health Records. *International Journal of Medical Informatics*. 112, (Apr. 2018), 59–67. DOI:https://doi.org/10.1016/j.ijmedinf.2018.01.007.

[21] Burns, M.N. et al. 2011. Harnessing context sensing to develop a mobile intervention for depression. *Journal of Medical Internet Research*. 13, 3 (Jul. 2011), e55. DOI:https://doi.org/10.2196/jmir.1838.

[22] Bü□enbender, P.C. et al. 2018. Ubiquitous computing applied to mental health: trends and research focus. (2018). DOI:https://doi.org/10.1145/3243082.3267456.

[23] Cai, L. et al. 2018. State affect recognition using smartphone sensing data. *Proceedings of the 2018 IEEE/ACM International Conference on Connected Health: Applications, Systems and Engineering Technologies* (New York, NY, USA, Sep. 2018), 120–125.

[24] Canzian, L. and Musolesi, M. 2015. Trajectories of depression: Unobtrusive monitoring of depressive states by means of smartphone mobility traces analysis. *UbiComp 2015 - Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing* (Sep. 2015), 1293–1304.

[25] Cao, J. et al. 2020. Tracking and Predicting Depressive Symptoms of Adolescents Using Smartphone-Based Self-Reports, Parental Evaluations, and Passive Phone Sensor Data: Development and Usability Study. *JMIR mental health*. 7, 1 (Jan. 2020), e14045. DOI:https://doi.org/10.2196/14045.

[26] Caudill, E.M. and Murphy, P.E. 2000. Consumer Online Privacy: Legal and Ethical Issues. *Journal of Public Policy & Marketing*. 19, 1 (Apr. 2000), 7–19. DOI:https://doi.org/10.1509/jppm.19.1.7.16951.

[27] Chen, Y. et al. 2017. Unpacking Happiness: Lessons from Smartphone Photography among College Students. *Proceedings - 2017 IEEE International Conference on Healthcare Informatics, ICHI 2017* (Sep. 2017), 429–438.

[28] Chervonsky, E. and Hunt, C. 2019. Emotion Regulation, Mental Health, and Social Wellbeing in a Young Adolescent Sample: A Concurrent and Longitudinal Investigation. *Emotion*. 19, 2 (Mar. 2019), 270–282. DOI:https://doi.org/10.1037/emo0000432.

[29] Ciman, M. and Wac, K. 2018. Individuals' Stress Assessment Using Human-Smartphone Interaction Analysis. *IEEE Transactions on Affective Computing*. 9, 1 (Jan. 2018), 51–65. DOI:https://doi.org/10.1109/TAFFC.2016.2592504.

[30] Constantinides, M. et al. 2018. Personalized versus generic mood prediction models in bipolar disorder. *UbiComp/ISWC 2018 - Adjunct Proceedings of the 2018 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2018 ACM International Symposium on Wearable Computers* (New York, New York, USA, Oct. 2018), 1700–1707.

[31] Core ML | Apple Developer Documentation: *https://developer.apple.com/documentation/coreml*. Accessed: 2020-11-02.

[32] Costs of Nine Common Mental Disorders: Implications for Curative and Preventive Psychiatry - PubMed: *https://pubmed.ncbi.nlm.nih.gov/17200596/*. Accessed: 2020-07-04.

[33] Dal researchers create app that monitors mental health by measuring user's emotions | CBC News: *https://www.cbc.ca/news/canada/nova-scotia/dalhousie-university-researcher-prosit-study-smartphone-app-mental-health-treatment-1.5653705*. Accessed: 2020-08-06.

[34] De, D. et al. 2015. Multimodal wearable sensing for fine-grained activity recognition in healthcare. *IEEE Internet Computing*. 19, 5 (Sep. 2015), 26–35.

DOI:https://doi.org/10.1109/MIC.2015.72.

[35] Degirmenci, K. 2020. Mobile users' information privacy concerns and the role of app permission requests. *International Journal of Information Management*. 50, (Feb. 2020), 261–272. DOI:https://doi.org/10.1016/j.ijinfomgt.2019.05.010.

[36] DeMasi, O. et al. 2016. Detecting change in depressive symptoms from daily wellbeing questions, personality, and activity. *2016 IEEE Wireless Health, WH 2016* (Dec. 2016), 22–29.

[37] Diagnostic and Statistical Manual of Mental Disorders (DSM-5®) - American Psychiatric Association - Google Books: *https://books.google.ca/books?hl=en&lr=&id=-JivBAAAQBAJ&oi=fnd&pg=PT18&dq=+Diagnostic+and+Statistical+Manual+of+Mental+Disorders&ots=ceWP-8IGB6&sig=IUgo889umXhjgIhU4xRMIY8tRoQ&redir_esc=y#v=onepage&q=Diagnostic and Statistical Manual of Mental Disorders&f=false*. Accessed: 2020-08-05.

[38] Doherty, G. et al. 2008. Technology in mental health. *Conference on Human Factors in Computing Systems - Proceedings* (New York, New York, USA, 2008), 3965–3968.

[39] Edirisooriya, Y.L. et al. 2019. Smartphone-based Approach to Enhance Mindfulness among Undergraduates with Stress. *2019 IEEE 14th International Conference on Industrial and Information Systems: Engineering for Innovations for Industry 4.0, ICIIS 2019 - Proceedings* (Dec. 2019), 464–469.

[40] Exposito, M. et al. 2018. Affective keys: Towards unobtrusive stress sensing of smartphone users. *MobileHCI 2018 - Beyond Mobile: The Next 20 Years - 20th International Conference on Human-Computer Interaction with Mobile Devices and Services, Conference Proceedings Adjunct* (New York, New York, USA, Sep. 2018), 139–145.

[41] Farhan, A.A. et al. 2016. Multi-view Bi-clustering to Identify Smartphone Sensing Features Indicative of Depression. *Proceedings - 2016 IEEE 1st International Conference on Connected Health: Applications, Systems and Engineering Technologies, CHASE 2016* (Aug. 2016), 264–273.

[42] Featured | Apple Developer Documentation: *https://developer.apple.com/documentation/*. Accessed: 2020-08-19.

[43] Gaggioli, A. et al. 2013. A mobile data collection platform for mental health research. *Personal and Ubiquitous Computing*. 17, 2 (Feb. 2013), 241–251. DOI:https://doi.org/10.1007/s00779-011-0465-2.

[44] Garcia-Ceja, E. et al. 2018. Mental health monitoring with multimodal sensing and machine learning: A survey. *Pervasive and Mobile Computing*. Elsevier B.V.

[45] Georgiades, K. et al. 2019. Six-Month Prevalence of Mental Disorders and Service Contacts among Children and Youth in Ontario: Evidence from the 2014 Ontario Child Health Study. *Canadian Journal of Psychiatry*. (2019). DOI:https://doi.org/10.1177/0706743719830024.

[46] Ghandeharioun, A. et al. 2019. EMMA: An Emotion-Aware Wellbeing Chatbot. *2019 8th International Conference on Affective Computing and Intelligent Interaction, ACII 2019* (Sep. 2019), 15–21.

[47] Ghosh, S. et al. 2019. EmoKey: An Emotion-aware Smartphone Keyboard for Mental Health Monitoring. *2019 11th International Conference on Communication Systems and Networks, COMSNETS 2019* (May 2019), 496–499.

[48] Ghosh, S. et al. 2017. TapSense: Combining Self-Report Patterns and Typing Characteristics for Smartphone based Emotion Detection. (2017). DOI:https://doi.org/10.1145/3098279.3098564.

[49] Glenn, T. and Monteith, S. 2014. New Measures of Mental State and Behavior Based on Data Collected From Sensors, Smartphones, and the Internet. *Current Psychiatry Reports*. Current Medicine Group LLC 1.

[50] Gong, J. et al. 2019. Understanding behavioral dynamics of social anxiety among college students through smartphone sensors. *Information Fusion*. 49, (Sep. 2019), 57–68. DOI:https://doi.org/10.1016/j.inffus.2018.09.002.

[51] González-Pérez Ignacio Miralles, A. et al. 2019. Technical Challenges to Deliver Sensor-

based Psy-chological Interventions Using Smartphones. (2019). DOI:https://doi.org/10.1145/3341162.3346271.

[52] Google AI Blog: Federated Learning: Collaborative Machine Learning without Centralized Training Data: *https://ai.googleblog.com/2017/04/federated-learning-collaborative.html*. Accessed: 2020-09-17.

[53] Gravenhorst, F. et al. 2015. Mobile phones as medical devices in mental disorder treatment: an overview. *Personal and Ubiquitous Computing*. 19, 2 (2015), 335–353. DOI:https://doi.org/10.1007/s00779-014-0829-5.

[54] Greenberg, P.E. and Birnbaum, H.G. 2005. The economic burden of depression in the US: Societal and patient perspectives. *Expert Opinion on Pharmacotherapy*. Taylor & Francis.

[55] Grünerbl, A. et al. 2015. Smartphone-based recognition of states and state changes in bipolar disorder patients. *IEEE Journal of Biomedical and Health Informatics*. 19, 1 (Jan. 2015), 140–148. DOI:https://doi.org/10.1109/JBHI.2014.2343154.

[56] Gu, J. et al. 2017. Privacy concerns for mobile app download: An elaboration likelihood model perspective. *Decision Support Systems*. 94, (Feb. 2017), 19–28. DOI:https://doi.org/10.1016/j.dss.2016.10.002.

[57] Hagiwara, N. et al. 2016. Validity of the Mind Monitoring System as a Mental Health Indicator. *Proceedings - 2016 IEEE 16th International Conference on Bioinformatics and Bioengineering, BIBE 2016* (Dec. 2016), 262–265.

[58] Hampel, F.R. 1974. The influence curve and its role in robust estimation. *Journal of the American Statistical Association*. 69, 346 (1974), 383–393. DOI:https://doi.org/10.1080/01621459.1974.10482962.

[59] Hard, A. et al. *FEDERATED LEARNING FOR MOBILE KEYBOARD PREDICTION*.

[60] Harmful Noise Levels | HealthLink BC: *https://www.healthlinkbc.ca/health-topics/tf4173*. Accessed: 2020-12-16.

[61] Henson, P. et al. 2020. Towards clinically actionable digital phenotyping targets in

schizophrenia. *npj Schizophrenia*. (2020). DOI:https://doi.org/10.1038/s41537-020-0100-1.

[62] Hilty, D. et al. 2020. A framework for competencies for the use of mobile technologies in psychiatry and medicine: Scoping review. *JMIR mHealth and uHealth*. JMIR Publications.

[63] Huang, K. et al. 2015. Monitoring Sleep and Detecting Irregular Nights through Unconstrained Smartphone Sensing. *2015 IEEE 12th Intl Conf on Ubiquitous Intelligence and Computing and 2015 IEEE 12th Intl Conf on Autonomic and Trusted Computing and 2015 IEEE 15th Intl Conf on Scalable Computing and Communications and Its Associated Workshops (UIC-ATC-ScalCom)* (2015), 36–45.

[64] Huang, Y. et al. Discovery of Behavioral Markers of Social Anxiety from Smartphone Sensor Data. DOI:https://doi.org/10.1145/3089341.3089343.

[65] Inal, Y. et al. 2020. Usability evaluations of mobile mental health technologies: Systematic review. *Journal of Medical Internet Research*. Journal of Medical Internet Research.

[66] Kang, H. and Shin, W. 2016. Do Smartphone Power Users Protect Mobile Privacy Better than Nonpower Users? Exploring Power Usage as a Factor in Mobile Privacy Protection and Disclosure. *Cyberpsychology, Behavior, and Social Networking*. 19, 3 (Mar. 2016), 179–185. DOI:https://doi.org/10.1089/cyber.2015.0340.

[67] Kendall, P.C. and Peterman, J.S. 2015. Evidence-Based Psychiatric Treatment CBT for Adolescents With Anxiety: Mature Yet Still Developing. *Am J Psychiatry*. (2015).

[68] Kirwan, M. et al. 2013. Smartphone apps for physical activity: A systematic review. *ASICS Conference / Journal of Science and Medicine in Sport*. 16, (2013), 39–58. DOI:https://doi.org/10.1016/j.jsams.2013.10.112.

[69] Konečn´y, J.K. et al. 2016. *Federated Optimization: Distributed Machine Learning for On-Device Intelligence*.

[70] Lanata, A. et al. 2015. Complexity index from a personalized wearable monitoring system for assessing remission in mental health. *IEEE Journal of Biomedical and Health*

Informatics. 19, 1 (Jan. 2015), 132–139. DOI:https://doi.org/10.1109/JBHI.2014.2360711.

[71] Lane, N.D. et al. 2010. A survey of mobile phone sensing. *IEEE Communications Magazine*. 48, 9 (Sep. 2010), 140–150. DOI:https://doi.org/10.1109/MCOM.2010.5560598.

[72] Layard, R. et al. 2012. How mental illness loses out in the NHS. (Jun. 2012).

[73] Lee, J. et al. 2018. Privacy-Preserving Patient Similarity Learning in a Federated Environment: Development and Analysis. *JMIR medical informatics*. 6, 2 (Apr. 2018), e20. DOI:https://doi.org/10.2196/medinform.7744.

[74] Lim, W.Y.B. et al. 2019. Federated Learning in Mobile Edge Networks: A Comprehensive Survey. (Sep. 2019).

[75] Liu, M. 2013. A study of mobile sensing using smartphones. *International Journal of Distributed Sensor Networks*. 2013, (Mar. 2013). DOI:https://doi.org/10.1155/2013/272916.

[76] Lonzetta, A. et al. 2018. Security Vulnerabilities in Bluetooth Technology as Used in IoT. *Journal of Sensor and Actuator Networks*. 7, 3 (Jul. 2018), 28. DOI:https://doi.org/10.3390/jsan7030028.

[77] Lu, H. et al. 2012. StressSense: Detecting stress in unconstrained acoustic environments using smartphones. *UbiComp'12 - Proceedings of the 2012 ACM Conference on Ubiquitous Computing* (2012), 351–360.

[78] Lu, J. et al. 2018. Joint Modeling of Heterogeneous Sensing Data for Depression Assessment via Multi-task Learning. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*. 2, 1 (Mar. 2018), 1–21. DOI:https://doi.org/10.1145/3191753.

[79] Ma, Y. et al. 2012. Daily mood assessment based on mobile phone sensing. *Proceedings - BSN 2012: 9th International Workshop on Wearable and Implantable Body Sensor Networks* (2012).

[80] Majumder, A.J.A. et al. 2019. SEmoD: A personalized emotion detection using a smart holistic embedded IoT system. *Proceedings - International Computer Software and Applications Conference* (Jul. 2019), 850–859.

[81] Malhotra, N.K. et al. 2004. Internet users' information privacy concerns (IUIPC): The construct, the scale, and a causal model. *Information Systems Research*. INFORMS Inst.for Operations Res.and the Management Sciences.

[82] Mankodiya, K. et al. 2013. Understanding user's emotional engagement to the contents on a smartphone display: Psychiatric prospective. *Proceedings - IEEE 10th International Conference on Ubiquitous Intelligence and Computing, UIC 2013 and IEEE 10th International Conference on Autonomic and Trusted Computing, ATC 2013* (2013), 631–637.

[83] Marzano, L. et al. 2015. The application of mHealth to mental health: Opportunities and challenges. *The Lancet Psychiatry*. Elsevier Ltd.

[84] Di Matteo, D. et al. 2018. Patient Willingness to Consent to Mobile Phone Data Collection for Mental Health Apps: Structured Questionnaire. *JMIR mental health*. 5, 3 (Aug. 2018), e56. DOI:https://doi.org/10.2196/mental.9539.

[85] Matthews, M. et al. 2014. Tracking mental well-being: Balancing rich sensing and patient needs. *Computer*. 47, 4 (2014), 36–43. DOI:https://doi.org/10.1109/MC.2014.107.

[86] McCarthy, M. and Spachos, P. 2017. Wellness assessment through environmental sensors and smartphones. *IEEE International Conference on Communications* (Jul. 2017).

[87] McLeod, K. et al. 2020. An Experimental Framework for Wellness Assessment Using the Internet of Things. *IEEE Internet Computing*. 24, 2 (Mar. 2020), 8–17. DOI:https://doi.org/10.1109/MIC.2020.2972817.

[88] McLeod, K. et al. 2020. Smartphone-Based Wellness Assessment Using Mobile Environmental Sensors. *IEEE Systems Journal*. (2020), 1–11. DOI:https://doi.org/10.1109/JSYST.2020.3004558.

[89] McMahan, B. and Ramage, D. 2017. Federated Learning : Collaborative Machine

Learning without centralized training data. *Post*. (2017). DOI:https://doi.org/10.1080/00173130902749999.

[90] McNamara, L. and Ngai, E. 2018. SADHealth: A Personal Mobile Sensing System for Seasonal Health Monitoring. *IEEE Systems Journal*. 12, 1 (Mar. 2018), 30–40. DOI:https://doi.org/10.1109/JSYST.2016.2521805.

[91] Mehrotra, A. et al. Towards Multi-modal Anticipatory Monitoring of Depressive States through the Analysis of Human-Smartphone Interaction. DOI:https://doi.org/10.1145/2968219.2968299.

[92] Mehrotra, A. and Musolesi, M. 2017. Designing effective movement digital biomarkers for unobtrusive emotional state mobile monitoring. *DigitalBiomarkers 2017 - Proceedings of the 1st Workshop on Digital Biomarkers, co-located with MobiSys 2017* (New York, New York, USA, Jun. 2017), 3–8.

[93] Mental Disorders & Conditions - DSM5: *https://psychcentral.com/disorders/*. Accessed: 2020-08-02.

[94] Mestry, M. et al. 2015. Identifying associations between smartphone usage and mental health during depression, anxiety and stress. *Proceedings - 2015 International Conference on Communication, Information and Computing Technology, ICCICT 2015* (Feb. 2015).

[95] Mishra, V. 2019. From Sensing to Intervention for Mental and Behavioral Health. (2019). DOI:https://doi.org/10.1145/3341162.3349304.

[96] Mohr, D.C. et al. 2017. Personal Sensing: Understanding Mental Health Using Ubiquitous Sensors and Machine Learning. *Annual Review of Clinical Psychology*. 13, 1 (May 2017), 23–47. DOI:https://doi.org/10.1146/annurev-clinpsy-032816-044949.

[97] Monteith, S. and Glenn, T. 2016. Automated Decision-Making and Big Data: Concerns for People With Mental Illness. *Current Psychiatry Reports*. Current Medicine Group LLC 1.

[98] Mood disorders - Symptoms and causes - Mayo Clinic: *https://www.mayoclinic.org/diseases-conditions/mood-disorders/symptoms-causes/syc-*

20365057. Accessed: 2020-09-05.

[99] Mushtaq, R. et al. 2014. Relationship between loneliness, Psychiatric disorders and physical health ? A review on the psychological aspects of loneliness. *Journal of Clinical and Diagnostic Research*. Journal of Clinical and Diagnostic Research.

[100] Narziev, N. et al. 2020. STDD: Short-term depression detection with passive sensing. *Sensors (Switzerland)*. 20, 5 (Mar. 2020). DOI:https://doi.org/10.3390/s20051396.

[101] NASA - This Week @ NASA, November 23, 2012: *https://www.nasa.gov/multimedia/podcasting/TWAN_11_23_12.html*. Accessed: 2020-07-04.

[102] Orji, R. et al. 2019. Socially-driven persuasive health intervention design: Competition, social comparison, and cooperation. *Health Informatics Journal*. 25, 4 (Dec. 2019), 1451–1484. DOI:https://doi.org/10.1177/1460458218766570.

[103] Orji, R. and Moffatt, K. 2018. Persuasive technology for health and wellness: State-of-the-art and emerging trends. *Health Informatics Journal*. 24, 1 (Mar. 2018), 66–91. DOI:https://doi.org/10.1177/1460458216650979.

[104] Osmani, V. et al. 2013. Monitoring activity of patients with bipolar disorder using smart phones. *ACM International Conference Proceeding Series* (2013), 85–92.

[105] Osmani, V. 2015. Smartphones in Mental Health: Detecting Depressive and Manic Episodes. *IEEE Pervasive Computing*. 14, 3 (Jul. 2015), 10–13. DOI:https://doi.org/10.1109/MPRV.2015.54.

[106] Palmius, N. et al. 2014. A multi-sensor monitoring system for objective mental health management in resource constrained environments. *IET Conference Publications* (2014).

[107] Palmius, N. et al. 2018. Group-personalized regression models for predicting mental health scores from objective mobile phone data streams: Observational study. *Journal of Medical Internet Research*. 20, 10 (Oct. 2018), e10194. DOI:https://doi.org/10.2196/10194.

[108] Place, S. et al. 2017. Behavioral indicators on a mobile sensing platform predict clinically validated psychiatric symptoms of mood and anxiety disorders. *Journal of Medical Internet Research*. 19, 3 (Mar. 2017), e75. DOI:https://doi.org/10.2196/jmir.6678.

[109] PROSIT: New app tracks your mental health based on how you use your phone: *https://www.bodyandsoul.com.au/mind-body/wellbeing/theres-now-an-app-that-tracks-your-mental-health-based-on-how-you-use-your-phone/news-story/4b837e48191ecc524cea49154355093d*. Accessed: 2020-08-06.

[110] Psychological Distress | Cardiac College: *https://www.healtheuniversity.ca/EN/CardiacCollege/Wellbeing/Stress_And_Sense_Of_Control/Pages/psychological-distress.aspx*. Accessed: 2020-09-05.

[111] Rabbi, M. et al. 2011. *Passive and In-situ Assessment of Mental and Physical Well-being using Mobile Sensors*.

[112] Rana, R. et al. 2016. Opportunistic and Context-Aware Affect Sensing on Smartphones. *IEEE Pervasive Computing*. 15, 2 (Apr. 2016), 60–69. DOI:https://doi.org/10.1109/MPRV.2016.36.

[113] Rohani, D.A. et al. 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 (Aug. 2018), e165. DOI:https://doi.org/10.2196/mhealth.9691.

[114] Sabatelli, M. et al. 2015. Correlation of significant places with self-reported state of bipolar disorder patients. *Proceedings of the 2014 4th International Conference on Wireless Mobile Communication and Healthcare - "Transforming Healthcare Through Innovations in Mobile and Wireless Technologies", MOBIHEALTH 2014* (Jan. 2015), 116–119.

[115] Saeb, S. et al. 2015. Mobile phone sensor correlates of depressive symptom severity in daily-life behavior: An exploratory study. *Journal of Medical Internet Research*. 17, 7 (Jul. 2015). DOI:https://doi.org/10.2196/jmir.4273.

[116] Saeb, S. et al. 2015. The relationship between clinical, momentary, and sensor-based assessment of depression. *Proceedings of the 2015 9th International Conference on Pervasive Computing Technologies for Healthcare, PervasiveHealth 2015* (Dec. 2015), 229–232.

[117] Sano, A. et al. 2018. Identifying objective physiological markers and modifiable behaviors for self-reported stress and mental health status using wearable sensors and mobile phones: Observational study. *Journal of Medical Internet Research*. 20, 6 (Jun. 2018), e210. DOI:https://doi.org/10.2196/jmir.9410.

[118] Saxena, S. et al. 2007. Resources for mental health: scarcity, inequity, and inefficiency. *Lancet*. Lancet.

[119] Servia-Rodríguez, S. et al. 2017. Mobile sensing at the service of mental well-being: A large-scale longitudinal study. *26th International World Wide Web Conference, WWW 2017* (2017), 103–112.

[120] Shiffman, S. et al. 2008. Ecological Momentary Assessment. *Annual Review of Clinical Psychology*. 4, 1 (Apr. 2008), 1–32. DOI:https://doi.org/10.1146/annurev.clinpsy.3.022806.091415.

[121] Silver, L. and Taylor, K. 2019. Smartphone Ownership Is Growing Rapidly Around the World, but Not Always Equally. *Pew Research Center*. (2019).

[122] Smith, H.J. et al. 1996. Information privacy: Measuring individuals' concerns about organizational practices. *MIS Quarterly: Management Information Systems*. 20, 2 (1996), 167–195. DOI:https://doi.org/10.2307/249477.

[123] Solove, D.J. 2006. A taxonomy of privacy. *University of Pennsylvania Law Review*.

[124] Spathis, D. et al. 2019. Passive mobile sensing and psychological traits for large scale mood prediction. *ACM International Conference Proceeding Series* (2019).

[125] Stangor, C. and Walinga, J. 2014. 13.1 Psychological Disorder: What Makes a Behaviour Abnormal? BCcampus.

[126] Stewart, K.A. and Segars, A.H. 2002. An empirical examination of the concern for information privacy instrument. *Information Systems Research*. 13, 1 (2002), 36–49. DOI:https://doi.org/10.1287/isre.13.1.36.97.

[127] Stigma and discrimination as a barrier to mental health service utilization in India | International Health | Oxford Academic: 2013. *https://academic.oup.com/inthealth/article/5/1/6/699270*. Accessed: 2020-08-06.

[128] Stress | Mental Health Foundation: *https://www.mentalhealth.org.uk/a-to-z/s/stress*. Accessed: 2020-09-05.

[129] Stress Symptoms, Signs, and Causes - HelpGuide.org: *https://www.helpguide.org/articles/stress/stress-symptoms-signs-and-causes.htm*. Accessed: 2020-09-05.

[130] StudentLife Dataset: *https://studentlife.cs.dartmouth.edu/dataset.html*. Accessed: 2020-08-06.

[131] Sun, Y. et al. 2019. Investigating Privacy and Information Disclosure Behavior in Social Electronic Commerce. *Sustainability*. 11, 12 (Jun. 2019), 3311. DOI:https://doi.org/10.3390/su11123311.

[132] Suver, C. et al. 2020. Bringing code to data: Do not forget governance. *Journal of Medical Internet Research*. JMIR Publications.

[133] TestFlight - Apple: *https://testflight.apple.com/*. Accessed: 2020-08-19.

[134] Torous, J. et al. 2018. Characterizing the clinical relevance of digital phenotyping data quality with applications to a cohort with schizophrenia. *npj Digital Medicine*. (2018). DOI:https://doi.org/10.1038/s41746-018-0022-8.

[135] Torous, J. et al. 2018. The Emerging Imperative for a Consensus Approach Toward the Rating and Clinical Recommendation of Mental Health Apps. *The Journal of nervous and mental disease*. (2018). DOI:https://doi.org/10.1097/NMD.0000000000000864.

[136] Tuso, P. 2014. Treatment progress indicator: application of a new assessment tool to

objectively monitor the therapeutic progress of patients with depression, anxiety, or behavioral health impairment. *The Permanente journal*. (2014). DOI:https://doi.org/10.7812/TPP/13-091.

[137] UNDESA 2014. Definition of youth. *United Nations Department of Economic and Social Affairs*. 2009 (2014), 1–7.

[138] Vagias, W. 2006. Likert-type scale response anchors. *Clemson International Institute for Tourism and …*. (2006). DOI:https://doi.org/10.1525/auk.2008.125.1.225.

[139] Valenza, G. et al. 2014. Advanced technology meets mental health: How smartphones, textile electronics, and signal processing can serve mental health monitoring, diagnosis, and treatment. *IEEE Pulse*. 5, 3 (2014), 56–59. DOI:https://doi.org/10.1109/MPUL.2014.2309582.

[140] Viron, M.J. and Stern, T.A. 2010. The impact of serious mental illness on health and healthcare. *Psychosomatics*. Elsevier Inc.

[141] Wahle, F. et al. 2016. Mobile Sensing and Support for People With Depression: A Pilot Trial in the Wild. *JMIR mHealth and uHealth*. 4, 3 (Sep. 2016), e111. DOI:https://doi.org/10.2196/mhealth.5960.

[142] Walkup, J.T. et al. 2008. Cognitive Behavioral Therapy, Sertraline, or a Combination in Childhood Anxiety. *New England Journal of Medicine*. (2008). DOI:https://doi.org/10.1056/nejmoa0804633.

[143] Wang, F. et al. 2019. Assessing mental stress based on smartphone sensing data: An empirical study. *Proceedings - 2019 IEEE SmartWorld, Ubiquitous Intelligence and Computing, Advanced and Trusted Computing, Scalable Computing and Communications, Internet of People and Smart City Innovation, SmartWorld/UIC/ATC/SCALCOM/IOP/SCI 2019* (Aug. 2019), 1031–1038.

[144] Wang, J. et al. 2018. Not just privacy: Improving performance of private deep learning in mobile cloud. *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (New York, NY, USA, Jul. 2018), 2407–2416.

[145] Wang, R. et al. 2016. CrossCheck: Toward passive sensing and detection of mental health changes in people with schizophrenia. *UbiComp 2016 - Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing* (New York, NY, USA, Sep. 2016), 886–897.

[146] Wang, R. et al. 2017. Predicting Symptom Trajectories of Schizophrenia using Mobile Sensing. *PACMInteract.Mob.WearableUbiquitousTechnol*. 1, 3 (2017), 110. DOI:https://doi.org/10.1145/3130976.

[147] Wang, R. et al. 2014. Studentlife: Assessing mental health, academic performance and behavioral trends of college students using smartphones. *UbiComp 2014 - Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing* (2014), 3–14.

[148] Wang, R. et al. 2018. Tracking Depression Dynamics in College Students Using Mobile Phone and Wearable Sensing. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol. 2, 1, Article*. 43, 1 (2018), 43. DOI:https://doi.org/10.1145/3191775.

[149] Wang, R. et al. 2015. Using opportunistic face logging from smartphone to infer mental health: Challenges and future directions. *UbiComp and ISWC 2015 - Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing and the Proceedings of the 2015 ACM International Symposium on Wearable Computers* (Sep. 2015), 683–692.

[150] Wang, W. et al. 2018. Sensing Behavioral Change over Time: Using Within-Person Variability Features from Mobile Sensing to Predict Personality Traits. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol*. 2, 141 (2018), 141. DOI:https://doi.org/10.1145/3264951.

[151] Wang, W. et al. 2020. Social Sensing: Assessing Social Functioning of Patients Living with Schizophrenia using Mobile Phone Sensing. *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (New York, NY, USA, Apr. 2020), 1–15.

[152] Weather API - OpenWeatherMap: *https://openweathermap.org/api*. Accessed: 2020-08-

19.

[153] What Are Anxiety Disorders? *https://www.psychiatry.org/patients-families/anxiety-disorders/what-are-anxiety-disorders*. Accessed: 2020-09-05.

[154] What Are Bipolar Disorders? *https://www.psychiatry.org/patients-families/bipolar-disorders/what-are-bipolar-disorders*. Accessed: 2020-09-05.

[155] What Are Personality Disorders? *https://www.psychiatry.org/patients-families/personality-disorders/what-are-personality-disorders*. Accessed: 2020-09-05.

[156] What Are Sleep Disorders? *https://www.psychiatry.org/patients-families/sleep-disorders/what-are-sleep-disorders*. Accessed: 2020-09-05.

[157] What is Agile? | Atlassian: *https://www.atlassian.com/agile*. Accessed: 2020-08-01.

[158] What Is Depression? *https://www.psychiatry.org/patients-families/depression/what-is-depression*. Accessed: 2020-09-05.

[159] What Is PTSD? *https://www.psychiatry.org/patients-families/ptsd/what-is-ptsd*. Accessed: 2020-09-05.

[160] What Is Schizophrenia? *https://www.psychiatry.org/patients-families/schizophrenia/what-is-schizophrenia*. Accessed: 2020-09-05.

[161] Whiteford, H.A. et al. 2013. Global burden of disease attributable to mental and substance use disorders: Findings from the Global Burden of Disease Study 2010. *The Lancet*. 382, 9904 (Nov. 2013), 1575–1586. DOI:https://doi.org/10.1016/S0140-6736(13)61611-6.

[162] Xu, H. et al. 2008. *Association for Information Systems AIS Electronic Library (AISeL) Examining the Formation of Individual' s Privacy Concerns: Toward an Integrative View Recommended Citation*.

[163] Xu, H. et al. 2012. Effects of individual self-protection, industry self-regulation, and government regulation on privacy concerns: A study of location-based services. *Information Systems Research*. 23, 4 (2012), 1342–1363. DOI:https://doi.org/10.1287/isre.1120.0416.

[164] Xu, H. et al. 2012. Measuring Mobile Users' Concerns for Information Privacy. *ICIS 2012 Proceedings*. (Dec. 2012).

[165] Xu, H. et al. 2011. The personalization privacy paradox: An exploratory study of decision making process for location-aware marketing. *Decision Support Systems*. 51, 1 (Apr. 2011), 42–52. DOI:https://doi.org/10.1016/j.dss.2010.11.017.

[166] Xu, H. and Teo, H.-H. 2004. *Association for Information Systems AIS Electronic Library (AISeL) Alleviating Consumers' Privacy Concerns in Location-Based Services: A Psychological Control Perspective Recommended Citation ALLEVIATING CONSUMERS' PRIVACY CONCERNS IN LOCATION-BASED SERVICES: A PSYCHOLOGICAL CONTROL PERSPECTIVE*.

[167] Yamamoto, N. et al. 2019. Physiological Stress Level Estimation Based on Smartphone Logs. *2018 11th International Conference on Mobile Computing and Ubiquitous Network, ICMU 2018* (Feb. 2019).

[168] Zulueta, J. et al. 2018. Predicting mood disturbance severity with mobile phone keystroke metadata: A biaffect digital phenotyping study. *Journal of Medical Internet Research*. 20, 7 (Jul. 2018), e241. DOI:https://doi.org/10.2196/jmir.9775.

[169] 2006. Chapter 31. Mental Disorders. *Disease Control Priorities in Developing Countries (2nd Edition)*.

[170] 2010. DSM IV-TR. *Handbook of Disease Burdens and Quality of Life Measures*.

[171] 2017. WHO | Depression and Other Common Mental Disorders. *WHO*. (2017).

APPENDIX A. **Privacy Perceptions scales**

Core construct	**Code**	**Measurement Item**	**Created by**	**Validated by**
Prior privacy experience (PPE)	PPE1	How often have you personally experienced incidents whereby your personal information was used by some company or e-commerce web site without your authorization?	Xu et al. 2012 [164]	Degirmenci 2020 [35]
	PPE2	How much have you heard or read during the last year about the use and potential misuse of the information collected from the Internet?		
	PPE3	How often have you personally been the victim of what you felt was an improper invasion of privacy?		
Computer anxiety (CA)	CA1	Computers are a real threat to privacy in this country.	Stewart and Segars 2002 [126]	Degirmenci 2020 [35]
	CA2	I am anxious and concerned about the pace of automation in the world.		
	CA3	I am sometimes frustrated by increasing automation in my home.		
Perceived control (PC)	PC1	How much control do you feel you have over your personal information that has been released?	Xu et al. 2012 [163]	Degirmenci 2020 [35]
	PC2	How much control do you feel you have over the amount of your personal information collected by mobile apps?		
	PC3	Overall, how much in control do you feel you have over your personal information provided to mobile apps?		
	PC4	How much control do you feel you have over who can get access to your personal information?		
	PC5	How much control do you feel you have over how your personal information is being used by mobile apps?		
App permission concerns (APC)	APC1	It would bother me when I am asked to accept these app permissions.	Smith et al. 1996 [122]	Degirmenci 2020 [35]
	APC2	I would think twice before accepting these app permissions.		
	APC3	It would bother me to accept these app permissions.		
Perceived surveillance (PS)		If I would accept these app permissions...	Xu et al. 2012 [164]	Degirmenci 2020 [35]
	PS1	I believe that my mobile device would be monitored at least part of the time.		
	PS2	I would be concerned that the app is collecting too much information about me.		
	PS3	I would be concerned that the app may monitor my activities on my mobile device.		
Perceived intrusion (PI)		If I would accept these app permissions...	Xu et al. 2008 [162]	Degirmenci 2020 [35]
	PI2	I feel that as a result, others would know about me more than I am comfortable with.		

	PI2	I believe that as a result, information about me that I consider private would be more readily available to others than I would want.		
	PI3	I feel that as a result, information about me would be out there that, if used, would invade my privacy.		
Secondary use of personal information (SU)		If I would accept these app permissions...	Smith et al. 1996 [122]	Degirmenci 2020 [35]
	SU1	I would be concerned that the app may use my personal information for other purposes without notifying me or getting my authorization.		
	SU2	I would be concerned that the app may use my information for other purposes.		
	SU3	I would be concerned that the app may share my personal information with other entities without getting my authorization.		
Intention to accept app permissions (INT)		Given these app permission requests, specify the extent to which you would accept these app permissions.	Malhotra et al. 2004 [81]	Degirmenci 2020 [35]
	INT1	unwilling–willing		
	INT2	unlikely–likely		
	INT3	not probable–probable		
	INT4	impossible -possible		
Privacy concern (PRC)	PRC1	While using my smartphone, I am concerned about the way companies or marketers collect and use my personal information.	Kang and Shin 2016 [66]	
	PRC2	While using my smartphone, I feel like I am being asked to disclose a large amount of personal information.		
	PRC3	While using my smartphone, I am concerned that a company will track me down.		
	PRC4	It bothers me when my personal information is gathered during my smartphone use.		
Trust in app developers and health care providers (TR)	TR1	I think health care providers must collect my personal information through a mobile app to provide me with a better service.	Kang and Shin 2016 [66]	
	TR2	I can count on most health app designers to protect my data privacy.		
Privacy protection behavior (PPB)	PPB1	PPB1: Read a privacy policy before downloading a mobile app to your smartphone.	Kang and Shin 2016 [66]	
	PPB2	PPB2: Decide not to install an app on your smartphone because you found out you would have to share personal information to use it.		
	PPB3	PPB3: Uninstall an app on your smartphone because you found out it was collecting personal information that you did not want to share.		
	PPB4	PPB4: Turn off the location tracking feature on your smartphone because you were worried about other people or companies		

		being able to access that information.		
	PPB5	PPB5: Check the privacy setting of your smartphone.		
Perceived benefit (PB)	PB1	PB1: Information disclosure can provide me with the convenience to instantly access the product information that I need.	Xu et al. 2011 [165]	Sun et al. 2019 [131]
	PB2	PB2: Information disclosure can provide me with personalized services.		
	PB3	PB3: Information disclosure can provide me with monetary rewards.		
	PB4	PB4: Information disclosure can provide me with entertainment.		
Behavioral Intention (BI)	BI1	BI1: I am likely to disclose my personal information to use mobile apps for health and wellbeing in the next 12 months.	Xu and Teo 2004 [166]	Xu et al. 2012 [164]
	BI2	BI2: I predict I would use mobile apps for health and wellbeing in the next 12 months.		
	BI3	BI3: I intend to use mobile apps for health and wellbeing in the next 12 months.		

APPENDIX B. **Permission to Use**

In presenting this thesis in partial fulfilment of the requirements for a Postgraduate degree from the Dalhousie University, I agree that the Libraries of this University may make it freely available for inspection. I further agree that permission for copying of this thesis in any manner, in whole or in part, for scholarly purposes may be granted by the professor or professors who supervised my thesis work or, in their absence, by the Head of the Department or the Dean of the College in which my thesis work was done. It is understood that any copying or publication or use of this thesis or parts thereof for financial gain shall not be allowed without my written permission. It is also understood that due recognition shall be given to me and to the Dalhousie University in any scholarly use which may be made of any material in my thesis.

Requests for permission to copy or to make other use of the material in this thesis in whole or part should be addressed to:

Head of the Faculty of Computer Science
6050 University Ave,
Dalhousie University,
Halifax, Nova Scotia,
Canada
B3H 1W5

APPENDIX C. **Thesis Approval Form**

DALHOUSIE University

Faculty of Graduate Studies

MASTER'S THESIS APPROVAL FORM

FILLABLE FIELDS **MUST** BE COMPLETED ELECTRONICALLY.

PLEASE ENSURE ALL INFORMATION APPEARS **EXACTLY THE SAME** ON THIS FORM AND THE THESIS TITLE PAGE. THE USE OF WHITEOUT IS NOT PERMITTED.

THIS FORM, SHOULD BE SUBMITTED TO THE FACULTY OF GRADUATE STUDIES UPON FINAL THESIS SUBMISSION.

STUDENT NAME: BANUCHITRA SURULIRAJ	STUDENT NUMBER: B00788664

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

Master of Computer Science

(Computer Science, Faculty of)

DEFENCE/APPROVAL DATE:

THESIS TITLE: A MOBILE SENSING APP FOR MENTAL HEALTH TO SUPPORT FEDERATED LEARNING

THE UNDERSIGNED HEREBY CERTIFY THEY HAVE READ AND RECOMMENDED TO THE FACULTY OF GRADUATE STUDIES FOR ACCEPTANCE THE ABOVE THESIS. PLEASE NOTE: ONLY EXAMINERS WITH A FACULTY OF GRADUATE STUDIES APPOINTMENT MAY VOTE ON THE OUTCOME OF AN EXAMINATION AND SIGN THIS FORM. THE CHAIR OF THE DEFENCE SHOULD NOT SIGN.

	NAME	SIGNATURE
EXAMINING COMMITTEE MEMBERS:	Dr. Rita Orji	
	Dr. Derek Reilly	
	Dr. Sandra Meier	

AS RESEARCH SUPERVISOR FOR THE STUDENT NAMED ABOVE, I CERTIFY I HAVE READ THE STUDENT'S DEFENDED DISSERTATION (TITLE ABOVE), HAVE APPROVED CHANGES REQUIRED BY THE EXAMINING COMMITTEE, AND RECOMMEND THE DISSERTATION TO FGS FOR ACCEPTANCE. I FURTHER CERTIFY THAT: 1) I HAVE READ AND UNDERSTAND THE DALHOUSIE THESIS LICENSE AGREEMENT THAT THE STUDENT WILL SIGN AND SUBMIT TO FGS UPON FINAL SUBMISSION; 2) I HAVE ENSURED THE STUDENT HAS COMPLIED WITH ALL REQUIRED ETHICAL GUIDELINES AS PER FGS REGULATION 10.1; 3) I HAVE ENSURED THE STUDENT HAS RECEIVED COPYRIGHT PERMISSIONS FOR PUBLISHED MANUSCRIPTS, PAPERS OR REPORTS AS PER FGS REGULATION 10.2.2.

	NAME	SIGNATURE
SUPERVISOR(S):	Dr. Rita Orji	

NOTE: THIS FORM SHOULD BE COMPLETED AND SIGNED BY THE COMMITTEE AT THE DEFENCE (EXCEPT THE SUPERVISOR, WHO MAY WITHHOLD SIGNATURE UNTIL REVISIONS ARE APPROVED, IF ANY). IF THERE ARE REMOTE PARTICIPANTS IN THE DEFENCE, THEY MAY SIGN A SEPARATE COPY. THE SIGNED COPIES TOGETHER WILL CONSTITUTE A SINGLE FULLY SIGNED DOCUMENT. THE FACULTY OF GRADUATE STUDIES REQUIRES ALL SIGNATURES TO BE **ORIGINAL** (NO SCANS OR FAXES).

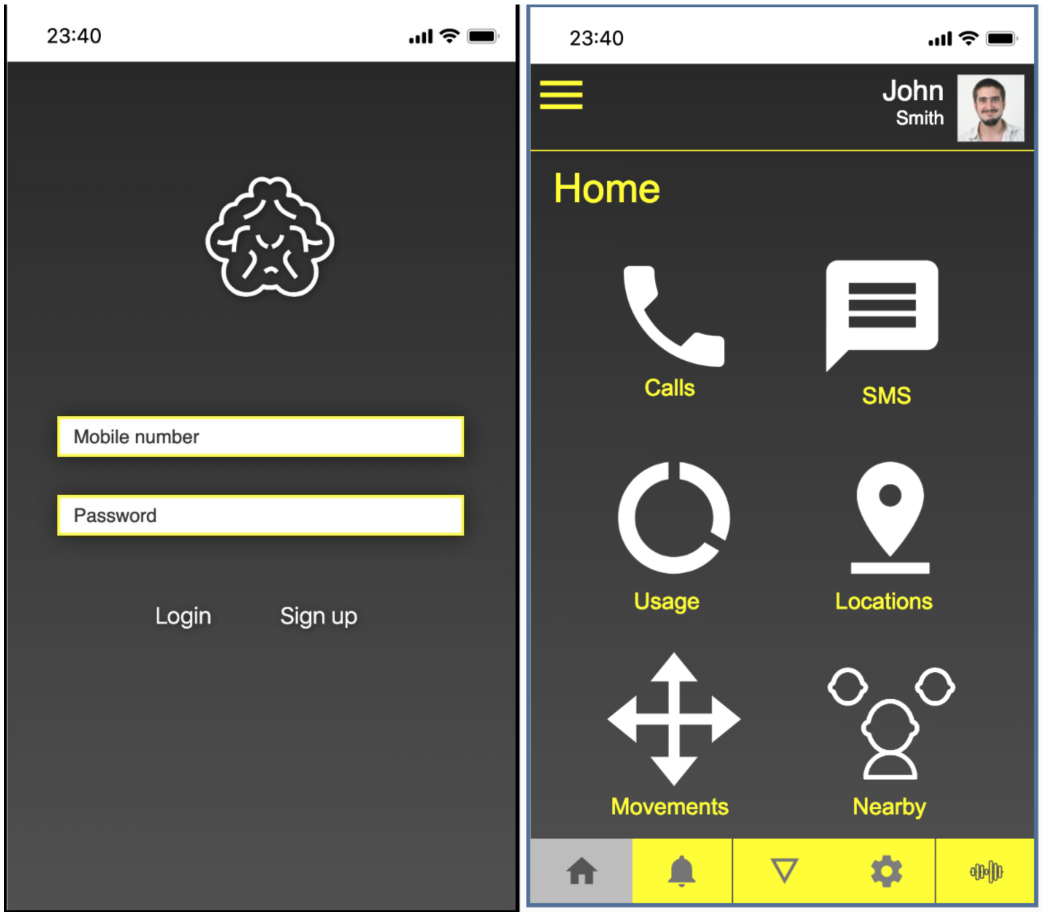
FACULTY OF GRADUATE STUDIES USE ONLY:

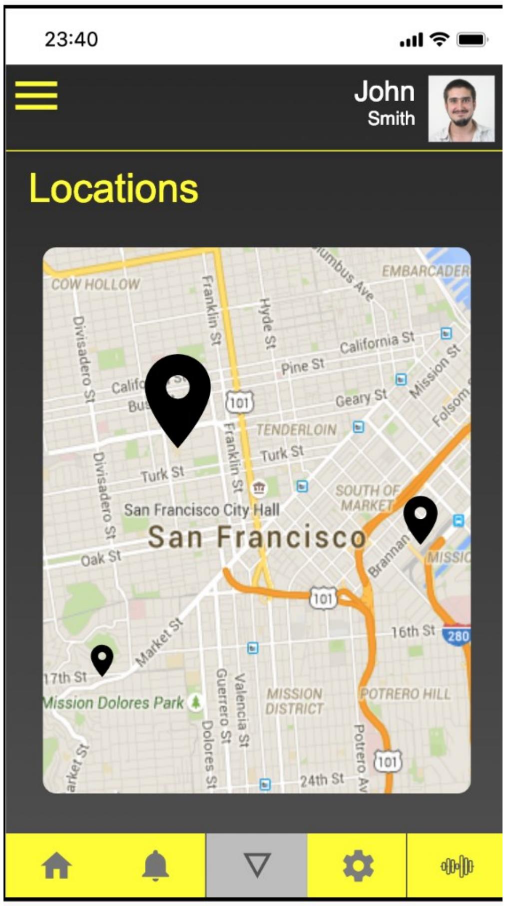
FACULTY OF GRADUATE STUDIES SIGNATURE

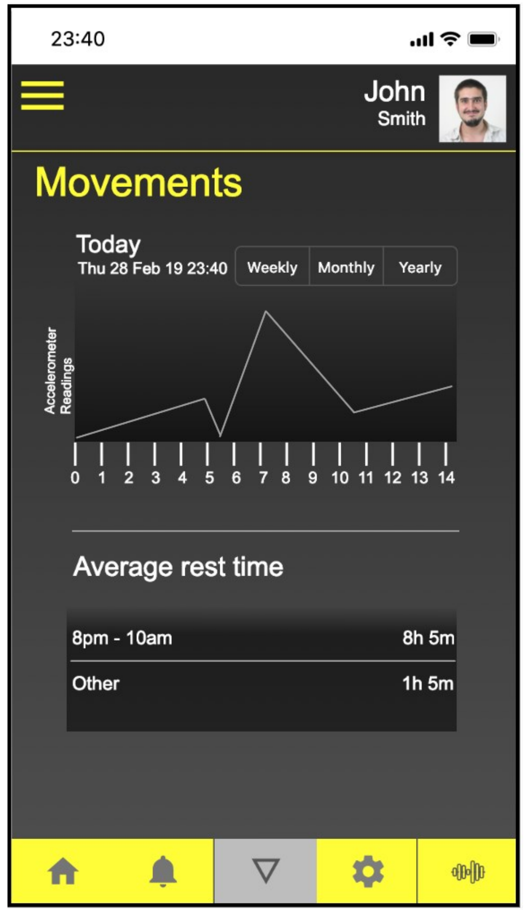
DATE THESIS ENTERED ON STUDENT RECORD

Revised May 2015 *Master's Thesis Approval Form*

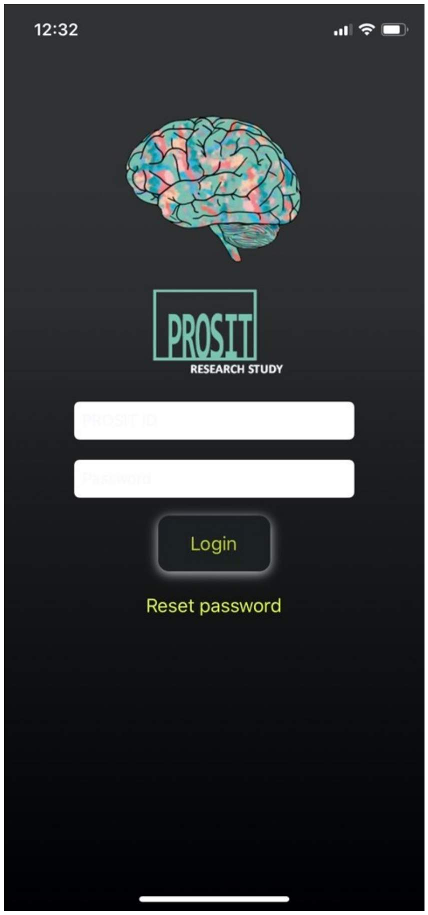
APPENDIX D. **Medium fidelity prototype PROSIT app**

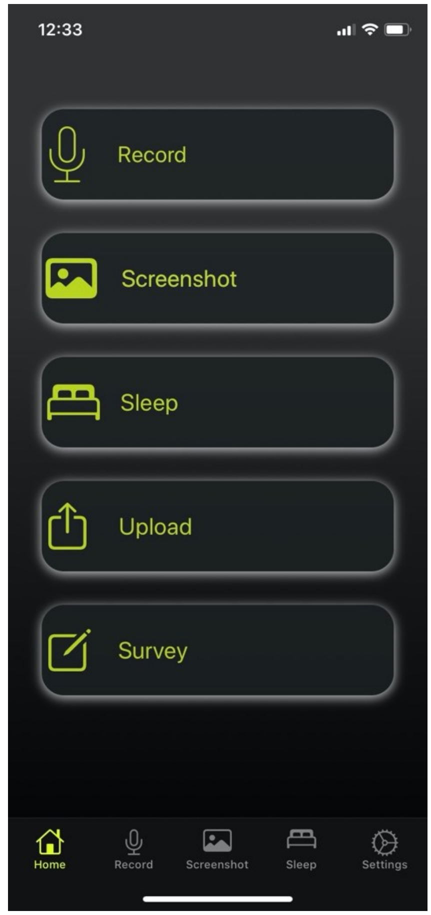


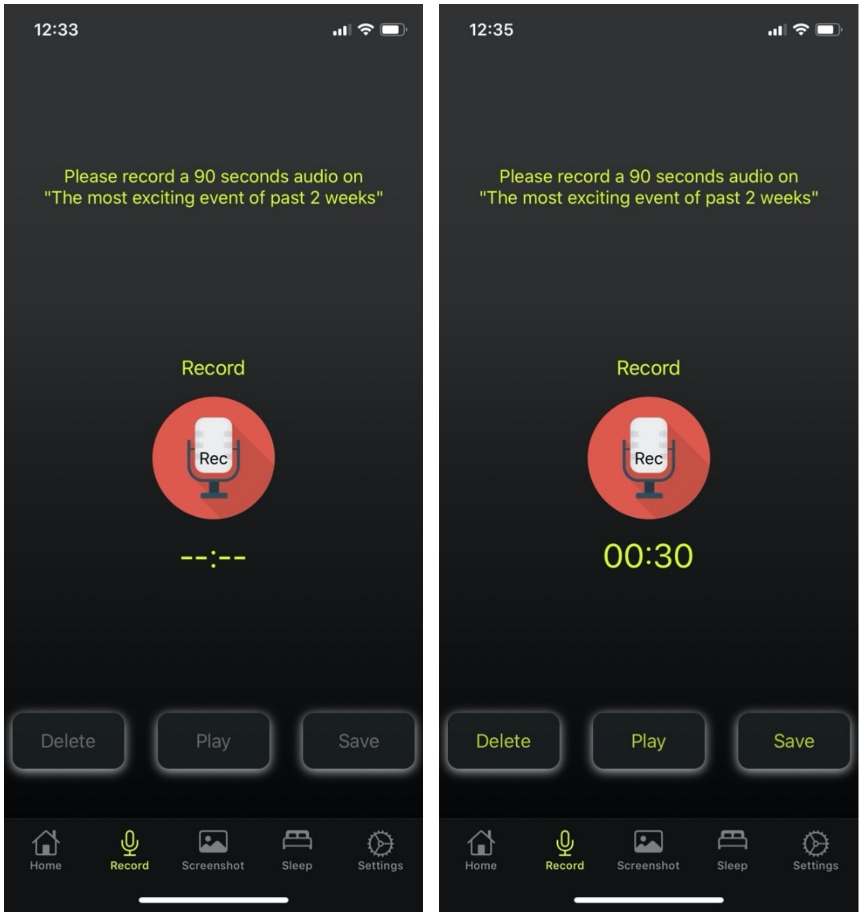

23:40
John
Smith
Locations
COW HOLLOW
EMBARCADERO
Franklin St
Hyde St
California St
Pine St
Geary St
Divisadero St
TENDERLOIN
Turk St
San Francisco City Hall
SOUTH OF MARKET
San Francisco
Oak St
Brannan St
16th St
17th St
Market St
Mission Dolores Park
MISSION DISTRICT
POTRERO HILL
Valencia St
Guerrero St
Dolores St
24th St
101
280

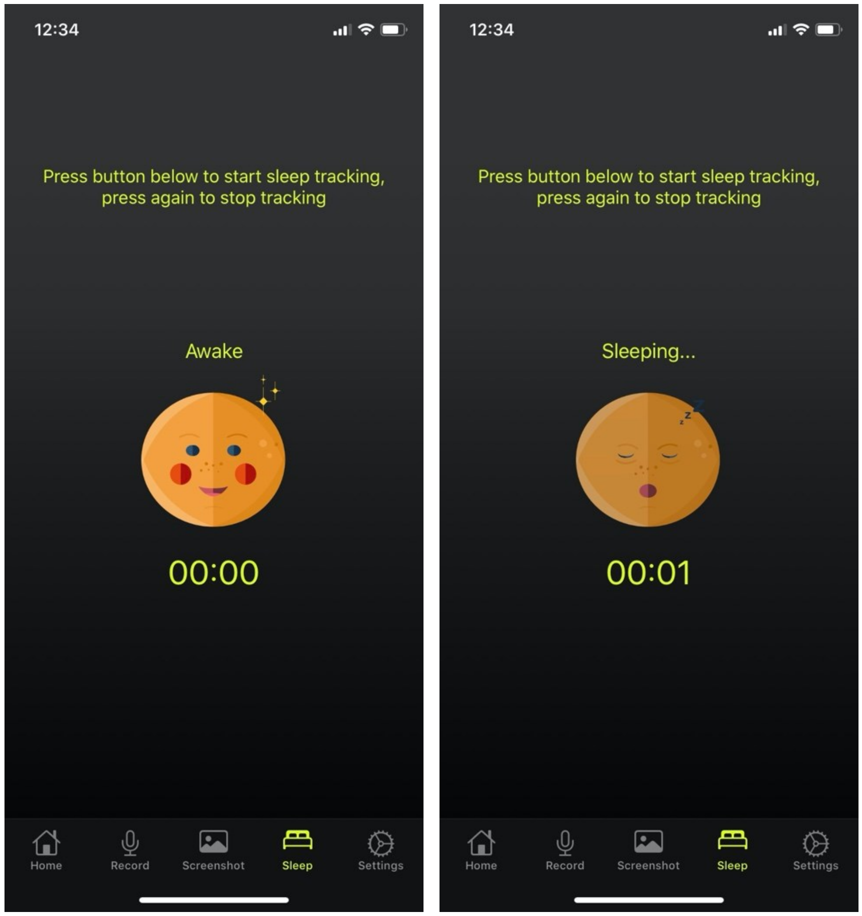

23:40
John
Smith
Movements
Today
Thu 28 Feb 19 23:40
Weekly
Monthly
Yearly
Accelerometer Readings
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14
Average rest time
8pm - 10am 8h 5m
Other 1h 5m

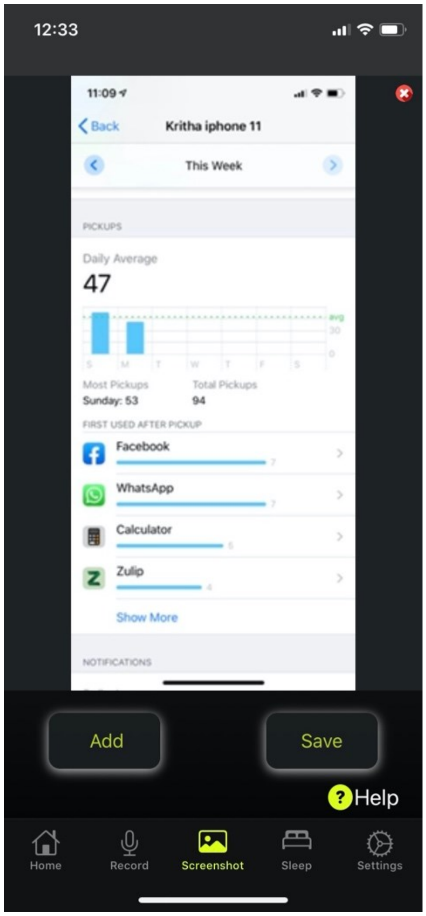
APPENDIX E. **Screenshots of PROSIT app**

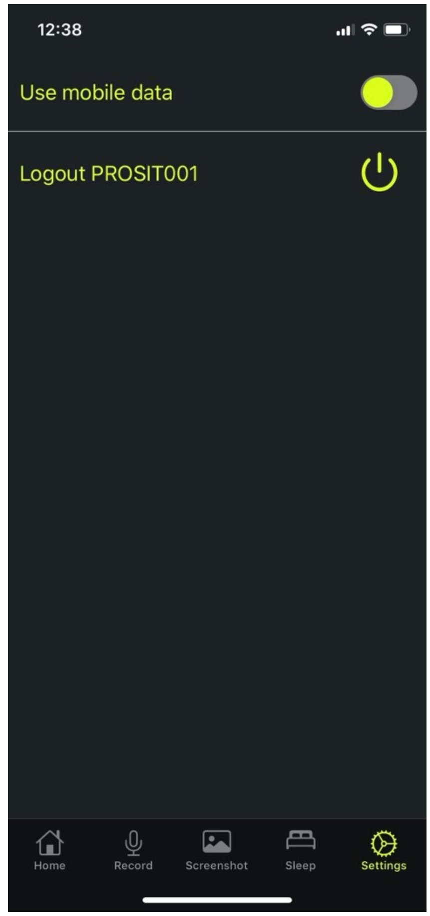





12:33
Please record a 90 seconds audio on
"The most exciting event of past 2 weeks"
Record
Rec
--:--
Delete
Play
Save
Home
Record
Screenshot
Sleep
Settings
12:35
Please record a 90 seconds audio on
"The most exciting event of past 2 weeks"
Record
Rec
00:30
Delete
Play
Save
Home
Record
Screenshot
Sleep
Settings


12:34
Press button below to start sleep tracking, press again to stop tracking
Awake
00:00
Home
Record
Screenshot
Sleep
Settings
12:34
Press button below to start sleep tracking, press again to stop tracking
Sleeping...
00:01
Home
Record
Screenshot
Sleep
Settings


12:33
11:09
Back
Kritha iphone 11
This Week
PICKUPS
Daily Average
47
Most Pickups
Sunday: 53
Total Pickups
94
FIRST USED AFTER PICKUP
Facebook
WhatsApp
Calculator
Zulip
Show More
NOTIFICATIONS
Add
Save
Help
Home
Record
Screenshot
Sleep
Settings


12:38
Use mobile data
Logout PROSIT001
Home
Record
Screenshot
Sleep
Settings

APPENDIX F. **Mturk Survey Questions**

Demographics

1. What is your age?

- 17 or less
- 18 to 24
- 25 to 34
- 35 to 44
- 45 and above

2. What is your gender?

- Male
- Female
- Other (Please specify)

3. What is your highest level of education?

- Less than High
- High school or
- College diploma
- Bachelor's degree
- Master's degree
- Doctorate degree
- Other (Please specify)

4. **Based on your cultural practices, which country you are associated with?**

[▾]

5. **Do you consider yourself a technologically knowledgeable person?**

Not at all	Not really	Somewhat	Moderate	Very much
○	○	○	○	○

6. **Have you ever used a health application to track or manage stress, low mood, depression, or anxiety?**

○ Yes

○ No

○ Not sure

7. **Have you ever used a mobile phone-based tracking application for health and wellness such as a fitness app, healthy eating, etc.?**

○ Yes

○ No

○ Not sure

8. **Which of the below COVID-19 issues have a significant impact on you? (Select all that apply)**

☐ Lockdown

☐ Quarantine

☐ Isolation

☐ Travel ban

☐ Financial issues

☐ Other (Please specify)

9. **To what extent does COVID-19 impact your overall health and wellbeing?**

Not at all	Slightly	Somewhat	Moderately	Very High

○	○	○	○	○

10. Do you use a mobile phone? If yes what type?

○ iPhone

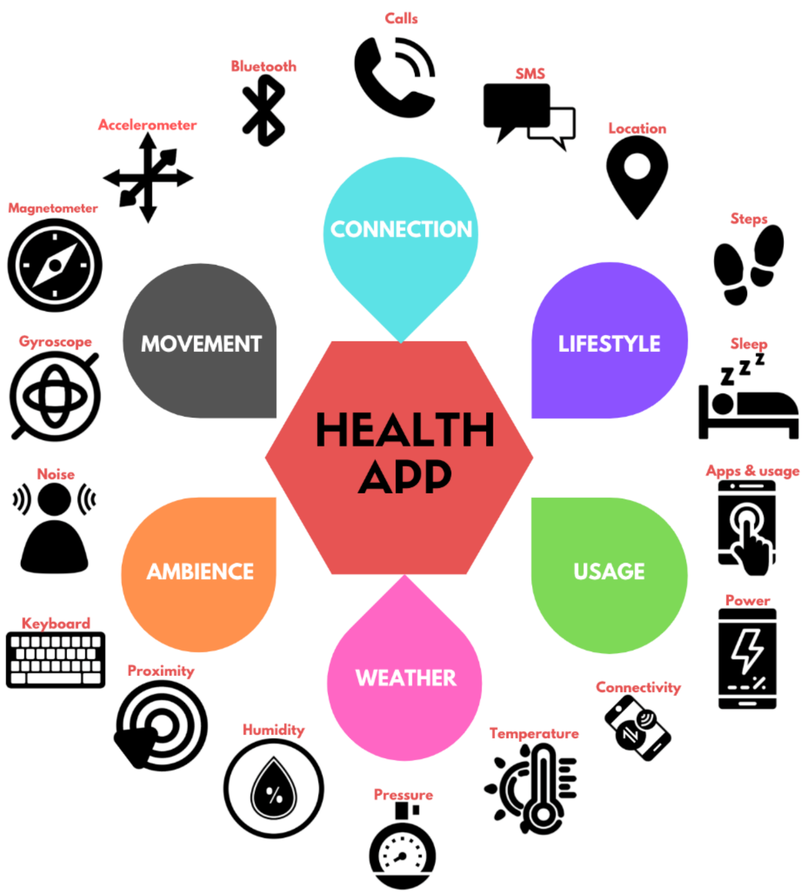
○ Android

○ Windows

○ I don't use a mobile

○ Other (Please specify)

11. Mobile phone-based tracking application for Health and Wellbeing - Overview



Location, Communication, Fitness, and Environment preferences

12. The following questions are to understand your opinion about tracking data from your mobile phone to help you improve your health and wellbeing. Answer the questions below assuming a mobile application has the capabilities explained. Rate your response on a scale between 1 to 5, 1 being lowest and 5 being highest, to what extent would you like the following features?

	1 – Not at all	2 - Slightly	3 - Moderately	4 – Very much	5 - Extremely
The app tracks your locations using GPS and notifies you when your patterns are unusual. Example. Too much time spent outside or not going out at all	◌	◌	◌	◌	◌
The app tracks your communication pattern and notifies you when your patterns are unusual. Example. Too much time spent on social media, or no communication at all	◌	◌	◌	◌	◌
The app tracks your mobile usage and predicts potential health issues, App automatically predicts and notify you. With no data sharing. If you are paying attention pick option 3	◌	◌	◌	◌	◌
The app tracks your health and sleep and notify you when your patterns are unusual. Example. Less/no sleep, missed exercise routines, less walk, or too much sleep, too much exercise.	◌	◌	◌	◌	◌
The app tracks your environment and notifies you when your patterns are unusual. Example. Too much noise, high temperature or less noise (always lonely), and freezing temperatures	◌	◌	◌	◌	◌

Data sharing preference

13. **If an app can track your mobile usage and predict potential health issues, what would you want the app to do with that information? Rate your preference level on the options below from 1 to 5. Where 1 being lowest and 5 being highest**

	1	2	3	4	5
App notify you and provide an option to share data with a doctor	○	○	○	○	○
App just notifies you. Data does not leave your device	○	○	○	○	○

Fine-grained tracking versus Summary tracking

14. **The following questions are to understand how comfortable you feel when data from your mobile phone is collected to predict your health issues. Answer the questions assuming a mobile application has the capabilities to collect data as explained below. Rate from 1 to 5 to state how much you are comfortable in the data collection feature, where 1 being lowest and 5 being highest**

Mobile usage data tracking - All the time						**Mobile usage data tracking - Hourly average**					
	1	2	3	4	5		1	2	3	4	5
Calls information: For every call, track data about call duration, type (incoming/outgoing). Does not collect the data about phone number or conversation	○	○	○	○	○	Calls information: Track hourly summary of the number of calls and duration	○	○	○	○	○
SMS information: For every text message, track data about length, type (sent/received). Does	○	○	○	○	○	SMS information: Track hourly count of the number of SMS sent/received,	○	○	○	○	○

not collect data on the content of messages						and count of typed keys					
Bluetooth information: Track data about names of connected and nearby devices. Does not collect data on content sent/received						Bluetooth information: Track hourly count of the number of Bluetooth devices nearby					
Location information: Track data about every street you visit. Does not collect data on the building address						Location information: Track hourly summary of location changes in distance					
Walking/running distance: Track data about the number of meters walked/ran and time of the walk/run						Walking/running distance: Track hourly sum of the number of meters walked					
Health information: Please select Moderately if you are reading carefully						Health information: Please select Extremely if you are reading carefully					
Light level: Track data about surrounding light levels all the time						Light level: Track hourly average of surrounding light level					
Noise level: Track data about the surrounding noise level all the time. Does not record audio						Noise level: Track hourly average of surrounding noise levels					
Weather: Track weather conditions like temperature, pressure, and humidity of the current place all the time						Weather information: Hourly average of temperature, pressure, and humidity of the environment					
Battery status: Track battery percentage all the time						Battery status: Track hourly average battery percentage					
Music: Track data about the name of the song for every played song						Music: Track hourly count of the number of songs played					
Movement: Track data about device movement,						Movement: Track hourly average of					

rotation, and magnetic field all the time						device movement, rotation, and magnetic field					
Screen time: Track data about app usage time for every app used on a daily basis	◌	◌	◌	◌	◌	Screen time: Track data about weekly average screen time	◌	◌	◌	◌	◌
Notification: Track data about the number of notifications received and the name of the app on a daily basis	◌	◌	◌	◌	◌	Notifications: Track weekly count of the number of notifications received and app names	◌	◌	◌	◌	◌
Device information: Track data about mobile phone model and operating system	◌	◌	◌	◌	◌	Device information: Track data about the mobile phone, choose option Extremely if you are reading carefully	◌	◌	◌	◌	◌
Pickups: Track data about count on the number of times the phone is picked up from rest, and first opened the app after pickup on a daily basis	◌	◌	◌	◌	◌	Pickups: Track the weekly average of the number of times the phone is picked from the rest	◌	◌	◌	◌	◌

15. To what extent are you knowledgeable about mobile tracking applications? Rate on a scale of 1-5 where 1 being poor and 5 being excellent

1- Not at all knowledge able	2 – Slightly knowledgeable	3 – Moderately knowledgeable	4 – Very much knowledgeable	5 – Extremely knowledgeable
◌	◌	◌	◌	◌

16. Mobile phone-based tracking applications can be used to determine your health and wellbeing?

1 – Strongly disagree	2 – Disagree	3 – Neither agree nor	4 - Agree	5 – Strongly agree

		disagree		
○	○	○	○	○

17. Did you ever receive treatment for mental health illness in the past/present?

○ Yes

○ No

○ Prefer not to

18. I have concerns about my mobile battery, my mobile battery drains fast

1 – Not at all	2 – Slightly	3 – Somewhat	4 - Moderately	5 – Very much
○	○	○	○	○

19. I'm concerned about my mobile internet (Mobile data) usage, and the associated cost incurred.

1 – Not at all	2 – Slightly	3 – Somewhat	4 - Moderately	5 – Very much
○	○	○	○	○

Prior privacy experience (PPE)

20. Answer the questions based on your prior experience and knowledge

	1 - Never	2 – Rarely	3. Sometimes	4 – Often	5 - Always
How often have you personally experienced incidents whereby your personal information was used by some company or eCommerce web	○	○	○	○	○

site without your authorization?					
How much have you heard or read during the last year about the use and potential misuse of the information collected from the Internet?	○	○	○	○	○
How often have you personally been the victim of what you felt was an improper invasion of privacy?	○	○	○	○	○

Computer anxiety (CA)

21. Rate your level of agreement to the following statements

	1 – Strongly Disagree	2 – Disagree	3 – Neither agree nor disagree	4 – Agree	5 – Strongly Agree
Computers and mobile phones are a real threat to privacy in this country.	○	○	○	○	○
I am anxious and concerned about the pace of automation in the world.	○	○	○	○	○
I am sometimes frustrated by increasing automation in my home.	○	○	○	○	○

Perceived control (PC)

22. Rate your perceptions of personal information security

	1 – Poor	2 – Fair	3 – Good	4 – Very good	5 - Excellent
How much control do you feel you have over your personal information that has been released online?	○	○	○	○	○

How much control do you feel you have over the amount of your personal information collected by mobile apps?	○	○	○	○	○
Overall, how much in control do you feel you have over your personal information provided to mobile apps?	○	○	○	○	○
How much control do you feel you have over who can get access to your personal information?	○	○	○	○	○
How much control do you feel you have over how your personal information is being used by mobile apps?	○	○	○	○	○

Consider a mobile application for health and wellbeing capable of tracking your mobile data such as lifestyle, usage, weather, ambiance, movement, and connection as detailed in previous sections, answer the below questions considering if you install the tracking mobile application and were to provide permissions.

App permission concerns (APC)

23. If I would install this app on my mobile device.

	1 – Not at all	2 – Slightly	3 – Moderately	4 – Very much	5 - Extremely
It would bother me when I am asked to accept the app permissions	○	○	○	○	○
It bothers me when I don't pay attention to the survey. choose option 5	○	○	○	○	○
I would think twice before accepting the app permissions	○	○	○	○	○
It would bother me to accept the app permissions.	○	○	○	○	○

Perceived surveillance (PS)

24. If I would accept the app permissions…

	1 – Not at all	2 – Slightly	3 – Moderately	4 – Very much	5 - Extremely
I believe that my mobile device would be monitored at least part of the time.	◌	◌	◌	◌	◌
I would be concerned that the app is collecting too much information about me	◌	◌	◌	◌	◌
I would be concerned that the app may monitor my activities on my mobile device.	◌	◌	◌	◌	◌

Perceived intrusion (PI)

25. If I would accept the app permissions…

	1 – Strongly Disagree	2 – Disagree	3 – Neither agree nor Disagree	4 – Agree	5 – Strongly Agree
I feel that as a result, others would know about me more than I am comfortable with.	◌	◌	◌	◌	◌
I believe that as a result, information about me that I consider private would be more readily available to others than I would want.	◌	◌	◌	◌	◌
I feel that as a result, information about me would be out there that, if used, would invade my privacy.	◌	◌	◌	◌	◌

Secondary use of personal information (SU)

26. If I would accept the app permissions…

	1- Not at all	2 – Slightly	3 – Moderately	4 – Very much	5 - Extremely
I would be concerned that the app may use my personal information for other purposes without notifying me or getting my authorization.	○	○	○	○	○
I would be concerned that the app may use my information for other purposes.	○	○	○	○	○
I would be concerned that the app may share my personal information with other entities without getting my authorization.	○	○	○	○	○

Intention to accept app permissions (INT)

27. Given the app permission requests, specify the extent to which you would accept these app permissions.

	1	2	3	4	5
Willingness	○	○	○	○	○
Likelihood	○	○	○	○	○
Probability	○	○	○	○	○
Possibility	○	○	○	○	○

Privacy concern (PRC)

28. Rate your level of concern on the following items

	1 – Not at all	2 – Slightly	3 – Moderately	4 – Very much	5 - Extremely
While using my smartphone, I am concerned about the way companies or marketers collect and use my personal information.	○	○	○	○	○
While using my smartphone, I feel like I am being asked to disclose a large amount of personal information.	○	○	○	○	○
While using my smartphone, I am concerned that a company will track me down.	○	○	○	○	○
It bothers me when my personal information is gathered during my smartphone use.	○	○	○	○	○

Trust in app developers and health care providers (TR)

29. Rate your perceptions on the following items

	1 – Not at all	2 – Slightly	3 – Moderately	4 – Very much	5 - Extremely
I think health care providers must collect my personal information through a mobile app to provide me with a better service.	○	○	○	○	○

I can count on most health app designers to protect my data privacy.					

Privacy protection behavior (PPB)

30. Rate how frequently you perform the actions listed below

	1 – Never	2 – Almost never	3 – Sometimes	4 – Often	5 - Always
Read a privacy policy before downloading a mobile app to your smartphone.					
Decide not to install an app on your smartphone because you found out you would have to share personal information to use it.					
Uninstall an app on your smartphone because you found out it was collecting personal information that you did not want to share.					
Turn off the location tracking feature on your smartphone because you were worried about other people or companies being able to access that information.					
Check the privacy setting of your smartphone.					

Perceived benefit (PB)

31. Rate your opinion on information disclosure on items below

	1 – Never	2 – Almost	3 – Sometimes	4 – Often	5 - Always

		never			
Information disclosure can provide me with the convenience to instantly access the product information that I need.	○	○	○	○	○
Information disclosure can provide me with personalized services.	○	○	○	○	○
Information disclosure can provide me with monetary rewards.	○	○	○	○	○
Information disclosure can provide me with entertainment.	○	○	○	○	○

Behavioral Intention (BI)

32. Rate your intentions on the items below

	1 – Extremely unlikely	2 – Unlikely	3 – Neutral	4 – Likely	5 – Extremely Likely
I am likely to disclose my personal information to use mobile apps for health and wellbeing in the next 12 months.	○	○	○	○	○
I predict I would use mobile apps for health and wellbeing in the next 12 months.	○	○	○	○	○
I intend to use mobile apps for health and wellbeing in the next 12 months.	○	○	○	○	○

Personality Scale

33. On a scale of 1 to 5, to what extent do you agree with the following statements.

I see myself as someone who:	Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree
is reserved.	○	○	○	○	○
is generally trusting.	○	○	○	○	○

tends to be lazy	○	○	○	○	○
is relaxed and handles stress well.	○	○	○	○	○
has few artistic interests.	○	○	○	○	○
If you read this, select 'Strongly disagree'.	○	○	○	○	○
is outgoing and sociable.	○	○	○	○	○
tends to find fault with others.	○	○	○	○	○
does a thorough job	○	○	○	○	○
gets nervous easily.	○	○	○	○	○
has an active imagination	○	○	○	○	○

34. A mobile phone-based tracking application for health and wellbeing will be helpful in the pandemic/crisis like COVID-19

1 – Not at all	2 – Slightly	3 – Moderately	4 – Very	5 - Extremely
○	○	○	○	○

35. **Do you have any comments or questions?**

36. **Please enter your Mturk ID**

APPENDIX G. **PROSIT Team consent**

APPENDIX H. **Research Ethics Board Approval Letter**

Health Sciences Research Ethics Board
Amendment Approval

February 11, 2020
Felwah Alqahtani
Computer Science\Computer Science

Dear Felwah,
REB #: 2019-4860
Project Title: Designing a Mobile Application for Promoting Mental Health and Well-being in youth
The Health Sciences Research Ethics Board has reviewed your amendment request and has approved this amendment request effective today, February 11, 2020.

Sincerely,

Dr. Lori Weeks, Chair

Amendment Request for an Approved Project

Date: February 1, 2020

To: Social Sciences and Humanities Research Ethics Board

REB File #: 2019-4860

Project Title: Designing a Mobile Application for Promoting Mental Health and Well-being

Principal Investigator: Felwah Alqahtani

Co-investigator: Banuchitra Suruliraj

Dear REB Members

In consultation with our Co-investigator and my supervisor (Dr. Rita Orji), we decided to include an additional survey to collect data related to privacy perceptions with respect to mental health applications. Based on that, we made changes to the research in the following sections.

Section	Content
SECTION 1. ADMINISTRATIVE INFORMATION	Added Co-investigator Banuchitra Suruliraj
2.1.1 Lay summary	Purpose of privacy perceptions survey and references
2.3.1 Recruitment	We will conduct a survey with 500 participants and collect their opinions
2.3.2 Recruitment plans and append recruitment instruments	The co-investigator will also use the same approach for recruiting participants. Appendix A3 shows the recruitment script for the privacy perceptions survey.
2.4.1 Informed consent process	Consent form information for Privacy perceptions survey (Appendix B3)
2.4.2 How participants will be given the opportunity to withdraw	If participants decide that they do not want to continue with the study, they may choose not to submit their questionnaires and there is no penalty attached. However, after submitting, the researcher will not be able to remove their response because surveys are completed anonymously.
2.5.2	We describe how privacy perception survey will be conducted in 2.5.2. The privacy perceptions survey will be conducted separately from the main survey and will be analyzed for design considerations in later stages of building a mobile application
2.5.3	Mturk recruitment and Compensation information. 0.50 $ US dollar will be given for the privacy perception survey. See 2.5.3 for more details
2.5.5	Role of Co-investigator: Co-investigator Banuchitra Suruliraj

	will be responsible for recruitment and for administering the privacy study and analyzing the results
Appendix A3	Recruitment notice for Privacy perceptions survey
Appendix B3	Informed Consent for the Privacy perceptions survey
Appendix F	Privacy perceptions questions

All changes were highlighted in in yellow of approved version. Because we are not addressing new research questions or taking the study in a new direction, we believe it is appropriate to request for an amendment. If any further material or information is required, please do not hesitate to contact me.

Regards,

Felwah Alqahtani

APPENDIX I. **PROSIT Dataset usage approval**

IWK Health Centre

DALHOUSIE UNIVERSITY
FACULTY OF MEDICINE
Department of Psychiatry

December 4, 2020

RE: Data Banuchitra Suruliraj's thesis

To who it might concern

In my function as the study lead of the PROSIT study, I do provide my consent for the usage of our pilot data in Banuchitra Suruliraj's thesis. The data includes the mobile sensing features of 18 youth sampled over 10 days.

The data is shared with Banuchitra Suruliraj and her supervisor Dr. Rita Orji under the assumptions that:

1. Data is pseudoanonymized (personally identifiable information fields within a data record are replaced by an artificial identifier).
2. There are no details provided about the diagnosis of the patients.
3. The data are destroyed after the thesis completion, the data or not shared or accessible to other parties.
4. The results are not getting published without consultation with me as the study lead.
5. All commercially reasonable safeguards are in place to ensure the confidentiality of the data and no data will be shared or accessible to any third parties without my written consent.

Sincerely yours,

Sandra Meier, PhD

Associate Professor,
Canada Research Chair in Developmental Psychopathology and Youth Mental Health Department of Psychiatry, Dalhousie University

Child & Adolescent Psychiatry • IWK Health Centre • 5850/5980 University Avenue • Halifax, NS B3K 6R8
Tel: 902.470.7720 • Fax: 902.470.7893

APPENDIX J. **Agile process timeline**

Month	Progress
November 2018	1. Literature review 2. High level requirements discussios: Calls, SMS, Application usage, Location, Battery and Screen 3. High level design diagram 4. Challenges in iOS Call, SMS and App usage data retrieval 5. iOS: Call and SMS data retrieval from call_history.db file in iOS operating system 6. Requirement: Include Missed call and cancelled call information 7. Jailbreaking option for iOS – this option is not feasible 8. Screentime feature is released in iOS
December 2018	1. Android: Implemented code for Call, SMS and App usage features 2. iOS: Implemented Call tracking in AppDelegate, SMS cannot be tracked. Dropped plan to retrieve from call_history.db 3. Analyzed "Freedom" and "Moment" applications for iOS screentime data 4. Discussion: iOS Enterprise developer and MDM options 5. VisionAPI to extract screentime data 6. Android: Implemented Android Call, SMS and App usage tracking 7. iOS: Explored TesseractOCR data extraction from Screentime screenshots
January 2019	1. iOS: Extracted screentime data using TesseractOCR 2. iOS: Implemented Accelerometer, Magnetometer and Gyroscope 3. Android: Implemented motion sensors Accelerometer, Magnetometer and Gyroscope 4. Plan: Buy a desktop computer and use it as a server in Mona

	Campbell building
February 2019	1. Bought and desktop computer 2. Setup MAMP server on the computer and attempted to access from outside the network 3. Faced firewall constraints, and Dalhousie CS Helpdesk denied public IP address for the desktop computer
March 2019	1. iOS: Implemented steps and sleep data extraction from HealthKit 2. Android: Implemented steps tracking 3. iOS: Implemented lock state tracking 4. iOS and Android: Implemented power state tracking
April 2019	1. Server: Requested Dalhousie IT Support for a server 2. Server: Requested IWK IT Support for a server 3. iOS: Implemented local SQLite database 4. Android: Implemented local SQLite database 5. Android: Data export option, view in local device
May 2019	1. Plan: Use webservice and https data transfer to upload data to server 2. Plan: Use MongoDB for data storage 3. Server: Implemented Firebase backend for login, and data upload for temporary usage
June 2019	1. iOS and Android: Implemented Reachability, ambient light, proximity, location and Bluetooth features 2. Proximity tracking not working in iOS – works only when the app is in use. Removed the feature 3. Bluetooth tracking – restricted in iOS 4. Crashes in Location tracking module, fixed the bug 5. Implemented App permissions module
July 2019	1. Dominik is hired for server setup and Android UI development

	2. Android: Knowledge transfer to Dominik 3. iOS: Collected UDID of 5 test iPhones, distributed the app for testing 4. Android: Apk distributed for internal testing in 7 devices
August 2019	1. iOS and Android: Design and develop user interface 2. Dominik implemented Apache server – it failed 3. Dominik completed server setup with Docker compose, nginx and MongoDB 4. Removed Firebase backend 5. Webservices are setup and active 6. iOS and Android: Implemented API and data upload 7. Issue: More than 100MB data generated per day 8. iOS and Android: Discussion and implementation of thresholds for all the sensing features
September 2019	1. Dominik implemented dashboard in server using Meteor 2. iOS: Image picker implemented for uploading Screentime screenshots 3. Android bug: App frequently killed by OS 4. Android bug fix: Implemented background execution mode in Android 5. Android: Notifications tracking 6. iOS and Android: Retrieve device information
October 2019	1. Android: Implemented Noise tracking 2. Android: Noise tracking thresholds set to 5 dB 3. iOS: Noise tracking in headset usage, later removed this feature 4. Trial code: Additional features Music tracking, keyboard tracking and Mood tracking – These features are not approved by ethics, and not included in the application.
November 2019	1. iOS and Android: Implemented Weather tracking

	2. iOS: App distribution using TestFlight 3. Poster presentation at Dalhousie Psychiatry research day – got feedback from multiple psychiatrists 4. Started participant recruitment

APPENDIX K. **Publications During My Master's Degree**

1. Nkwo, M., Suruliraj, B., & Orji, R. (2020). Public perception of mental illness: Opportunity for community-based collaborative intervention. Conference on Human Factors in Computing Systems - Proceedings. https://doi.org/10.1145/3334480.3383023
2. Nkwo, M., Suruliraj, B., Orji, R., & Ugah, J. (2020). Socially-oriented persuasive strategies and sustainable behavior change: Implications for designing for environmental sustainability. CEUR Workshop Proceedings.
3. Suruliraj, B., Nkwo, M., & Orji, R. (2020). Persuasive Mobile Apps for Sustainable Waste Management: A Systematic Review. Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 12064 LNCS, 182–194. https://doi.org/10.1007/978-3-030-45712-9_14
4. Suruliraj, B., Olagunju, T., Nkwo, M., & Orji, R. (2020). Bota: A personalized persuasive mobile app for sustainable waste management. CEUR Workshop Proceedings.

APPENDIX L. **Publications Related to My Thesis**

Title	**Venue**	**Status**
Comfortability with the Passive Collection of Smartphone Data for Monitoring of Mental Health: An Online Survey	Computers in Human Behavior Reports	Submitted
Mobile Sensing Apps and Self-management of Mental Health during the COVID-19 Pandemic– an Online Survey	JMIR Formative Research	Submitted
Mobile sensing applications for Mental health – Systematic review	International Journal of Mobile Human-Computer Interaction	Planned
PROSIT, a comprehensive mobile application for Mental health monitoring: Design, development and pilot study	Personal and Ubiquitous Computing	Planned
Federated learning framework for Mobile sensing applications for Mental health monitoring	Pervasive and Mobile Computing	Planned