Title	Loneliness detection using technology
Authors	Qirtas, Malik Muhammad
Publication date	2024
Original Citation	Qirtas, M. M. 2024. Loneliness detection using technology. PhD Thesis, University College Cork.
Type of publication	Doctoral thesis
Rights	© 2024, Malik Muhammad Qirtas. - https://creativecommons.org/licenses/by-nc-nd/4.0/
Download date	2026-09-25 04:07:10
Item downloaded from	https://hdl.handle.net/10468/16955

Loneliness Detection Using Technology

Malik Muhammad Qirtas

Thesis submitted for the degree of Doctor of Philosophy

NATIONAL UNIVERSITY OF IRELAND, CORK

College of Science, Engineering and Food Science

School of Computer Science & Information Technology
Centre for Research Training in Advanced Networks for Sustainable Societies

August 2024

Head of School: Professor Utz Roedig

Supervisor: Professor Dirk Pesch
Professor Eleanor Bantry White

Contents

List of Figures

List of Tables

I, Malik Muhammad Qirtas, certify that this thesis is my own work and has not been submitted for another degree at University College Cork or elsewhere.

Malik Muhammad Qirtas

Acknowledgements

First, I thank my supervisor, Prof. Dirk Pesch, for his great support and guidance throughout my Ph.D. journey. His supervision, patience, and understanding over the past four years have been invaluable. This thesis would not have been possible without his mentorship and dedication to my success. Having Dirk as my supervisor was a privilege and an honour.

I am deeply grateful to my co-supervisor, Prof. Eleanor Bantry White, whose interdisciplinary approach was really helpful in broadening the scope and impact of my research. Her input, mentorship, and ideas for bridging computer science and social sciences concepts greatly improved my research. This was very important in shaping how my final thesis turned out.

I want to say thanks to Dr. Evi Zafeiridi, whose support throughout my research, particularly during our data collection efforts, was invaluable. Her willingness to lend an ear and offer help with daily challenges made a significant difference in my experience.

To my fellow colleagues and friends, your support, both direct and indirect, has been a source of encouragement and inspiration throughout this journey. I am grateful for the collaborative and supportive environment you all helped to create.

I want to acknowledge the important role of my mentor and former manager, Sadia Ashraf, in my intellectual growth. Her guidance and support have been instrumental in my career as a software developer and researcher. I am deeply grateful for the lasting impact she has had on both my professional and personal growth.

I would like to express my deepest gratitude to my parents for their unwavering love, support, and encouragement throughout my academic journey. Their belief in my abilities and constant guidance have been my foundation, providing me with the strength to pursue my goals even in the most challenging moments. This accomplishment would not have been possible without their sacrifices and endless support.

Finally, I am grateful to the ADVANCE CRT for this opportunity. I want to say thanks to the program coordinators Luis Gómez De Membrillera and James Duggan, who always helped me with every aspect and whenever I needed any help. Cheers!

Malik Muhammad Qirtas
Cork, August 2024

Abstract

Over the last decade, especially during the COVID-19 pandemic, loneliness has emerged as a global issue. Although loneliness is a common issue that most people experience at some stage in their lives, it can have harmful effects on both physical and mental health if this condition becomes chronic. To avoid the long-term effects, detecting it in the early stages is critical, which would lead to timely intervention and treatment. Modern smartphones and wearables are equipped with a wide range of sensors that can provide a vast amount of information on people's daily lives and behaviours in real time, which can help detect early signs of loneliness. Researchers collect extensive data on a user's daily life and behavioural patterns, including social activities, mobility patterns, communication, and activity data through sensors like accelerometers, heart rate sensors, microphones, and GPS and Bluetooth connectivity acting as proximity sensors. All these capabilities of smartphones and wearables have made them a powerful tool for monitoring users' health and well-being passively. Many studies have used passive sensing for loneliness detection in recent years, but these methods have severe limitations. The main issues are related to relying on generic models, and neglecting the dynamic, multifaceted nature of loneliness. This thesis made several contributions to address these critical gaps in the literature.

First, we presented an approach for personalized loneliness detection through the behavioural grouping of passive sensing data. We used unsupervised machine learning for this and then trained customized models for each identified group that outperformed generic models for loneliness detection. Building upon this work, we addressed its limitation of static approach for loneliness detection. So, we proposed an approach for dynamic loneliness detection that can adjust to users' changing behaviours. We also analyzed how users' behaviours changed throughout the study period and how these changes are linked with their loneliness levels. This contribution helps detect loneliness in its early stages and highlights the importance of recognizing individual differences in how loneliness manifests.

Previous research on loneliness with passive sensing data often analyzed loneliness as a single concept. This thesis explored the different types of loneliness and their behavioural manifestations. Mainly, we investigated how behavioural patterns can distinguish between social and emotional loneliness and what is the predictive power of these digital biomarkers to detect the specific type of loneliness in order to develop more targeted interventions. Research shows that loneliness and depression often co-occur, but their behavioural overlap wasn't well understood in terms of behavioural patterns extracted through passive sensing. We investigated the complex relationship between these two conditions and identified both overlapping and distinct behavioural markers captured through passive sensing along with the predictive power of these markers. This work helps us understand how they interact with each other, and this could lead to improved detection and intervention strategies for individuals experiencing these conditions.

Chapter 1

Introduction

1.1 Overview

Loneliness is a global problem that affects people of all backgrounds. Recent studies, like the European Loneliness Survey, highlight just how common loneliness has become and show its status as a global concern [Joi23]. The implications of loneliness extend far beyond the subjective experience of isolation. It is increasingly recognized for its serious physical and mental health consequences [HLMA19]. Loneliness is not just about feeling alone; it can be detrimental to one's health if it becomes chronic. Loneliness has been associated with a higher likelihood of developing other mental health and physical health issues such as cardiovascular diseases and stroke [HLMA19]. Research has indicated that an absence of healthy social networks may contribute to an increased risk of these issues. Additionally, loneliness can disrupt hormones and compromise the body's immune system, making it more difficult to maintain good health [UUHL99]. It can also affect stress levels, blood pressure, and sleep quality [CEB+00], all of which can contribute to more severe health problems or even reduce one's lifespan. Moreover, chronic loneliness can increase mental health problems such as severe depression, suicidal thoughts, difficulty sleeping, and a general feeling of sadness [CHE+06]. The social and economic costs associated with loneliness are substantial. It results in a significant burden on healthcare systems due to increased demand for mental health services and physical health complications arising from prolonged loneliness [MLC+20]. The resulting effect on productivity and societal well-being further highlights the need for effective strategies to address this challenge. Loneliness is a big problem that got even worse during the COVID-19 pandemic. We need to understand loneliness better in order to develop effective and precise detection methods, which would be a step toward mitigating the negative health and societal impacts of this growing problem.

1.2 What is Loneliness?

Humans inherently seek social connections, and maintaining positive social interactions is essential for our well-being [BL17]. Perlman defined loneliness as a state of emotional suffering resulting from a discrepancy between the desired and actual social interactions [PP81]. Loneliness is not just limited to being physically alone; one might experience feelings of loneliness even in the presence of others. Unfortunately, loneliness sometimes bears a social stigma which indicates that an individual is unable to establish meaningful relationships [Rok12]. It has long been a worldwide problem, and it became even more noticeable during the COVID-19 pandemic. Before the pandemic, surveys showed that approximately one in four individuals reported feeling lonely. During the pandemic, that increased to almost two in every five people feeling lonely [SJD+23]. This shows just how much a big crisis can affect how lonely people feel all over the world.

It is important to understand that loneliness and being alone are not the same thing. Being alone is simply a physical state, while loneliness is a complex psychological experience. One can be alone and not feel lonely, and the opposite is also true: one can feel extremely lonely even when with others. This highlights that loneliness originates from the subjective quality of our social connections, not just their quantity. Furthermore, it is important to distinguish loneliness from social isolation. Social isolation describes a lack of social contact and interaction. Someone might be socially isolated without feeling lonely, or they might feel lonely even while maintaining a social network.

To further understand the nuances of loneliness, it is essential to distinguish between different types, which we will be referring to in this thesis (Chapter 6).

- Social Loneliness: This refers to the perceived lack of a social network, such as belonging to a group, having friends, or feeling like a part of a community [RCRY84]. Individuals experiencing social loneliness often feel disconnected from others and lack a sense of belonging.
- Emotional Loneliness: This happens when an individual feels the absence of close, intimate, meaningful relationships [Vau88]. Even those with a broad social network can experience emotional loneliness if they lack deep personal connections.

- Dispositional Loneliness: This refers to a stable personality trait where an individual consistently experiences feelings of loneliness, regardless of their external social environment or actual level of social interaction. It is often linked to factors such as low self-esteem, social anxiety, and negative self-perception [ROKvdM+12].

1.3 The Global Impact of Loneliness: A Growing Health Crisis

Loneliness has become a global issue that manifests across various demographics and geographical regions. This section provides a detailed overview of the global prevalence of loneliness and discusses its widespread nature.

The prevalence of loneliness is not limited to specific age groups, genders, or socio-economic classes. According to the European Loneliness Survey 2023, around 13% of respondents reported feeling lonely most or all of the time, while 36% reported feeling lonely at least some of the time [Joi23]. In a 2021 survey, younger adults, especially those aged between 18 and 34 years, reported higher levels of loneliness compared to the older age groups, with 16.8% feeling lonely "most of the time" or "always" [BBDC+21]. The COVID-19 pandemic has further increased the issue of loneliness, particularly among young adults and women [AOASAR+23]. During the pandemic, loneliness prevalence rates surged, with up to 20% of respondents in certain studies reporting feelings of loneliness [AOASAR+23]. Younger generations, specifically adolescents and young adults, report higher levels of loneliness compared to older adults. Some studies reported that 20–48% of adolescents and young adults experience severe levels of loneliness. However with older populations, more than 60% report never feeling lonely, and fewer than 10% report severe loneliness [BK20].

Furthermore, socio-economic factors play a significant role in the experience of loneliness. Income level, employment status, and educational history are important factors that could have a direct relationship with loneliness levels. Studies suggested that lower income and education levels, as well as unemployment are associated with increased chances of loneliness [HSSW20]. The prevalence rates of loneliness also vary considerably across Europe, with countries like the Netherlands reporting as low as 6% and others such as Bulgaria, Hungary, and Croatia experiencing rates over 16%. This variation shows the complexity of cultural, economic, and social factors in loneliness experiences. Socio-economic issues have a broader impact that goes beyond individual circumstances. A study reported that country-level income inequality correlates with the prevalence of loneliness, particularly among older adults [TMSA+22]. This suggests that broader socio-economic dynamics, including income distribution and relative deprivation, can substantially impact the experience of loneliness. This extensive prevalence shows the importance of understanding loneliness in its various aspects and contexts to address and mitigate its impact worldwide effectively.

The impact of loneliness on health and mental well-being is significant and multifaceted. Many studies highlight its close relationship with both physical and mental health issues which are given below.

- Physical Health Risks: Studies have found a significant association between loneliness and many physical health issues. Loneliness has been linked with cardiovascular diseases, inflammation, and even an increased risk of functional decline and death [PMG+20].

- Mental Health Implications: Loneliness significantly impacts mental health. It is associated with higher rates of depression, anxiety, and poorer quality of life [LCHS+20, HLHS+22]. The impact is higher in older adults, where loneliness is a known risk factor for dementia and other health issues [SANA20, VAIF+22]. Similarly, the association between loneliness and various forms of anxiety, including social and generalized anxiety, shows moderate to strong connections [MPTK+20]. Individuals experiencing loneliness have significantly higher odds of suffering from anxiety disorders. Longitudinal studies reinforce this, which identify loneliness as a predictive factor for developing anxiety disorders [MND+19]. This evidence shows the role of loneliness as a risk factor for mental health conditions and highlights the urgent need for early detection and intervention. Studies across various age groups have consistently shown a strong correlation between loneliness and the severity of depressive symptoms [IBY21, QZWP23, KAGH20, WPA+22].

- Impact on Immune Function and Sleep: Loneliness can affect immune function and increase susceptibility to getting infected more quickly [ZJ21]. It is also associated with disrupted sleep, which results in affecting immune function, glucose regulation, and cardiovascular risk [GWR+20, WBK+20]. Other sleep issues like insomnia are also commonly found in lonely individuals, which can further increase health issues [MJFS21].

1.4 The Challenge of Early Loneliness Detection and The Importance of Addressing Loneliness in Younger Populations

Monitoring early indicators of loneliness can enable us to intervene timely in order to help those who are suffering from this issue. Traditionally, clinicians rely on in-person consultations and patient self-disclosure to identify these signs, which are not always the most reliable methods and come with several challenges. First, there is a risk of underreporting, where individuals might not fully disclose their true feelings. It could be due to several reasons, such as a desire to appear socially content or, in some cases, to exaggerate their actual feelings to gain sympathy. Additionally, self-reports may not always be accurate. People might lack awareness of their own loneliness or find it difficult to express their emotions correctly. Misinterpretations of questionnaire language can also lead to incorrect responses. Another major challenge is recall bias. These traditional methods often capture feelings from a specific time frame, which may not accurately represent an individual's experience at other times. Memory biases can also alter the reporting of past events or emotions which can lead to further inaccuracies. These challenges highlight the need for more objective and continuous monitoring methods. The adoption of advanced technologies in smartphones and wearable devices has opened up new possibilities for addressing these issues.

Early detection of loneliness is critical for the success of interventions. Timely detection can lead to more effective ways to manage loneliness. This can prevent the conversion of loneliness into more severe mental health issues. However, early loneliness detection comes with significant challenges, particularly in populations that may not readily express or recognize their loneliness. One of the primary challenges in early detection is the subjective nature of loneliness. Loneliness is a personal experience that varies greatly among individuals, which makes it difficult to establish universal criteria for its detection. This subjective nature of loneliness can lead to underreporting, mainly in groups or cultures where a stigma has been associated with loneliness for different reasons. Another challenge is the lack of standardized screening tools for loneliness. While there are scales and questionnaires designed to measure loneliness, such as the University of California, Los Angeles (UCLA) Loneliness Scale [RPC80], their usage is not common in clinical or community settings [oSoBS20]. This lack of standardized screening means that loneliness is often overlooked in routine health assessments, mainly in primary care settings where such identification could be most beneficial.

Some specific populations are at higher risk of loneliness, and at the same time they could be more challenging to reach. For example, older adults, especially those in long-term care facilities may experience loneliness but lack the means or willingness to express it [ÖTFM18]. Similarly, individuals with chronic illnesses are at increased risk of social isolation and loneliness and they may not have enough opportunities to communicate their feelings with others [RLKS06]. Additionally, individuals with disabilities face unique challenges that can increase their vulnerability to loneliness. This includes physical barriers to social participation, stigma, and limited access to support networks [GZPA23]. It also significantly affects marginalized groups, such as the LGBTQ+ community. Research has consistently shown that LGBTQ+ individuals experience higher rates of loneliness compared to their heterosexual and cisgender counterparts [BWA+23, PIDWD20]. Similarly, the younger population despite being highly connected through digital platforms, may not recognize or admit their feelings of loneliness. This population is often masked by superficial online interactions that do not have any deeper meanings in their lives [SH21]. Younger people and college students are also a challenging demographic for the early detection of loneliness. The complexity is in the unique social dynamics, psychological development stages, and the recent impact of global crises, such as the COVID-19 pandemic on this age group [McA21].

The prevalence of depression, anxiety, and loneliness among college students has reached high levels, and it is worsened by different factors like the pandemic, political unrest, racism, and inequality. An extensive survey of around 33,000 college students found that a majority of two-thirds had difficulties with loneliness and a sense of isolation, showing the significant impact of these combined stress factors [McA21]. The COVID-19 pandemic significantly affected students' social lives and relationships, which highlights the impact of social crises [BSF20].

The early detection of loneliness in younger populations is important but at the same time very challenging due to the unique stressors and transitional phases associated with this population. This population group is increasingly at risk of loneliness. Multiple factors like life transitions, academic pressures, and evolving social dynamics contribute to this increased vulnerability [Ars21]. The pervasive use of technology among young people makes technological interventions more relevant and effective for this group. Their familiarity and comfort with digital devices give us a unique opportunity to reach them where they are most accessible.

The normalization of health monitoring in the young population, particularly through wearable devices and smartphone apps gives a good scope for passive sensing and early detection of loneliness indicators.

1.5 Passive Sensing for Loneliness Detection

The use of passive sensing methods is increasingly recognized as an important tool in the early and timely detection of different conditions including loneliness. Passive sensing technology, which includes the use of smartphones, wearables, and environmental sensors provides a non-intrusive way to monitor behavioural and physiological indicators that may be indicative of loneliness signs.

One of the primary advantages of passive sensing is its ability to collect continuous and objective data unobtrusively without the user's active involvement. This can track patterns in communication, social interactions, physical activity, and even sleep, which are often disruptive in individuals experiencing loneliness [FSRB22]. For example, changes in communication patterns, such as reduced frequency of calls or messages, could be indicative of social withdrawal which is a common symptom of loneliness. Similarly, the changes in usual sleep patterns, which we can detect through wearable devices are also indicative of some issue with mental well-being. Integrating passive sensing data with machine learning algorithms can lead to the early detection of signs of loneliness. These algorithms can analyze complex datasets to identify patterns that may not be immediately apparent to individuals or healthcare providers. For instance, a decrease in social activity or a change in daily routine detected through passive sensing can trigger alerts and prompt early intervention.

Moreover, technology can assist in bridging the gap in loneliness interventions. This is especially useful for populations that are hard to reach or reluctant to seek help due to the associated stigma with loneliness. Digital platforms and telehealth services can provide access to support and resources for individuals who might not otherwise engage with traditional healthcare systems. This is especially relevant for younger populations who are more used to technology and for older adults who may face mobility or accessibility issues. Studies exploring technological solutions for detecting loneliness highlight the potential of these tools in improving intervention strategies. For example, research into smartphone app usage patterns has shown promise in identifying loneliness indicators which shows that digital behaviour can be a valuable source of information for detecting loneliness [CH18]. Recent developments in smart technology, especially smartphones and fitness trackers, have significantly improved health monitoring capabilities [MSR+20].

Smartphone sensors like accelerometers and others can be used to distinguish between stationary and active users. In addition, the GPS sensor can be used to track social patterns, including the frequency and locations of visits, which may provide details about the level of social activity or isolation. Smartphones also have text and call logs, which can provide data about the frequency of our interpersonal interactions. This may serve as a significant indicator of our social connections. Bluetooth and WIFI logs can help us understand the proximity and distinct encounters with others. They can also be indicative of where a person spends the majority of their time based on the logs of the networks these sensors are connected to.

The type and usage of mobile apps are also important factors. For example, higher use of social media apps may indicate our intention to connect with people, but an increased duration spent on gaming applications may imply our preference for solitude. Similarly, the frequency with which one opens their phone may provide information about digital usage behaviours. Researchers can develop a more comprehensive understanding of human activities and behaviours by integrating the data collected from various sensors. All this information can be helpful with different analyses, and we can detect early loneliness signs which would result in timely intervention.

Similarly, fitness bands provide continuous physiological data. They measure step count, which is an effective measure of daily activity levels. These levels can be indicative of behavioural trends that are often linked to feelings of loneliness. Fitness bands also keep track of our heart rate. If the heart rate is different from usual, it could be an indicator of stress or feelings of strong emotions which could be linked with loneliness. Furthermore, detailed sleep pattern analysis from these wearables can provide behavioural markers of overall well-being.

An important part of validating data collected from passive sensors is using a reliable way to measure what we are actually trying to detect (loneliness). This is where clinical scales come in. Clinical scales are standardized tools developed by researchers to measure specific psychological condition or behavioural traits. These scales help provide a ground truth which is a reliable baseline measure that can be compared against the data collected from sensors. By using both passive data collection and clinical scales, researchers can check if what they are seeing in the sensor data actually corresponds to real feelings of loneliness that people report. The University of California, Los Angeles (UCLA) Loneliness Scale [RPC80], serves as the standard for assessing the subjective experience of loneliness. It is the most widely used scale in the general population and clinical settings for assessing the frequency and severity of subjective loneliness in an individual. This self-report questionnaire consists of a set of questions that respondents evaluate based on the frequency with which they experience certain feelings. The validity and reliability of this scale have undergone rigorous testing in various demographics and healthcare settings, and it has been established as an important tool for measuring loneliness by researchers.

In studies using passive sensing for loneliness detection, the UCLA scale is administered at various stages throughout the data collection process. The responses serve as labels that establish a connection between behavioural patterns obtained from sensors and self-reported feelings of loneliness. This direct input from participants about their perceived loneliness provides a benchmark against which the efficacy of machine learning models for loneliness prediction can be measured since these responses act as labels while training models. It ensures that the algorithm's interpretation of data aligns with the actual loneliness feelings of the users, which will improve the predictive power of the model. Through the periodic usage of the UCLA scale, researchers can also track changes over time. This provides information about the dynamic nature of loneliness and the impact of various interventions on managing loneliness.

Passive sensing has been extensively investigated in the context of various mental and physical health conditions. Many studies have focused on a broad range of conditions, including depression, stress, mood disorders, schizophrenia, sleep disturbances, cardiac health, and other chronic conditions [DVC+19, FMG+21, KL19, SNČS21, SQK21, DALW+22, AL06]. These studies have used a wide range of features extracted from sensor data, such as physical activity, sleep patterns, communication logs (call and SMS), and social interactions (via GPS or Bluetooth data), in combination with machine learning algorithms to predict different conditions. For example, [DVC+19] applied these features to detect depression, while [KL19] focused on detecting stress through the same features. These features are also relevant to loneliness detection but require different interpretive frameworks. For example, while high phone usage is usually associated with stress and depressive symptoms in previous studies, in the context of loneliness, it might reflect an attempt to seek social connection through digital means. Similarly, reduced physical activity may indicate withdrawal from social interactions which is a key behavioural marker of loneliness. In this thesis, we build upon these established features but focus on how they relate specifically to loneliness. By analyzing features like call and SMS logs, Bluetooth and WIFI proximity data, and location changes, we aim to detect social withdrawal or lack of social connectivity, which are key indicators of loneliness. While these features have been successfully applied to detect different mental health conditions, their contextual relevance to loneliness in a personalised manner has been underexplored. This thesis aims to fill that gap by applying these features to a more personalized and dynamic approach for loneliness detection.

While a number of studies have explored the potential of passive sensing for loneliness detection, there has been an absence of systematic reviews or surveys that synthesize the existing research on this specific topic. To bridge this gap, we performed a pioneering scoping review dedicated exclusively to the application of passive sensing in the context of loneliness and social isolation detection. This review is the first in the field to provide an in-depth look at the current capabilities of these technologies for loneliness detection and highlight the limitations that need to be addressed in future work.

The existing research shows multiple approaches that combine machine learning algorithms with passive sensing data to detect indicators of loneliness in diverse populations. These methods range from the analysis of communication logs, capturing call and text message data, to the application of machine learning techniques on sensor data that tracks physical activity and social interactions through accelerometers and GPS. Ambient sensors within smart home environments have been used to monitor routine domestic activities, while wearable devices have been used to collect physiological data for further analysis. Machine learning methods, especially those using predictive analytics and pattern recognition, have played an important role in analyzing this data. This body of work collectively validates the efficacy of passive sensing and machine learning as promising tools for the early detection of loneliness. However, this comes with a number of limitations and potential issues that set a direction for researchers to improve these technological approaches further.

This brief section provided the groundwork for the detailed exploration of passive sensing for loneliness detection in Chapter 2, where we will look into the methodology, findings, and implications of our scoping review with a comprehensive discussion of our research contributions.

1.5.1 Limitations of Existing Research

Existing studies in the field have made significant contributions to using passive sensing for loneliness detection. However, these approaches come with several limitations. Our work aims to address some of these limitations and improve upon existing methods. Below, we present the primary challenges that need to be addressed:

1.5.1.1 Privacy and Ethical Concerns

One of the key challenges is the issue of privacy. Most of the studies used a traditional machine learning model training process in which all data from multiple participants is collected on a single server, which trains the model on the whole data set at once. The collection of this sensitive data in a single location (central server) presents an important risk to the privacy of users, as it could expose them to unauthorized access and data breaches. When studies gather all their data in one place, it can make people worried about their privacy, especially with the sensitive topic of mental health. People might not want to be part of research that could expose their personal information. While it is important to address this issue, another important distinction should be made between the training and inference phases of machine learning. The goal is to avoid centralizing data during the model training process but once models are trained, they can be applied locally on personal devices for inference. It is really important to fix this data centralisation problem as when people feel sure that their private information is safe, they are more likely to use these health-tracking tools.

1.5.1.2 Generalized Models and Lack of Personalization

Existing methods for detecting loneliness with passive sensing often rely on generic models trained on the entire dataset. Given the subjective nature of loneliness, behavioural patterns can differ among people in relation to their feelings of loneliness. These generic models may not accurately capture the nuances of individual differences in behavioural patterns and this can lead to less precise detection. By identifying smaller groups of individuals with similar behavioural patterns, we can develop more targeted models that better reflect these individual differences. This approach can help us to uncover how loneliness manifests in specific groups, which can be overlooked when analyzing the entire population as a whole. Furthermore, monitoring the evolution of these behavioural groups over time can provide important information about how loneliness changes and adapts to individual experiences and circumstances which can ultimately lead to more timely and effective interventions.

1.5.1.3 Static Approaches and Lack of Real-Time Monitoring

Most studies in the domain of loneliness detection focus on a static analysis of passive sensing data, often collecting data over a fixed period and then analyzing it retrospectively. This limitation restricts the capacity of such systems to quickly detect and intervene in instances of loneliness that may arise suddenly as a consequence of life events. This highlights the need for a methodology that not only monitors but also adjusts to the evolving behavioural trends linked to loneliness. Such an approach has the capacity to provide more timely and personalized interventions that may stop the conversion of intermittent loneliness into a chronic condition like depression.

1.5.1.4 Multidimensional Nature of Loneliness

There is no existing work on examining the various dimensions of loneliness through the lens of passive sensing data. As we discussed earlier, loneliness is a multifaceted phenomenon that manifests in multiple forms. One of them is emotional loneliness, which is caused by the absence of a close emotional attachment, and social loneliness, which can be linked to perceived deficiencies in one's social network [SP22]. Acknowledging and analyzing the different facets of loneliness is important for a deeper understanding of the specific type of loneliness an individual is experiencing. This understanding is a key step towards developing personalized interventions that directly address the particular aspects and nuances of loneliness.

1.5.1.5 Relationship Between Loneliness and Depressions

Understanding the link between loneliness and depression is critical because they often occur together and can each make the other worse [WC23]. No study has yet explored how loneliness relates to depression and vice versa by using passive sensing and behavioural data. Since chronic loneliness can develop into more severe mental health issues, it's important to broaden our exploration of how behavioural data can be used to explore the relationship between loneliness and depression. By doing so, we can develop better strategies to help those who suffer from both depression and loneliness.

The following subsection presents the main contributions that are aimed at addressing some of the above issues and limitations.

1.6 Thesis Contributions

The key contributions of this thesis are summarized below:

1. The first contribution of this thesis presents a comprehensive scoping review. This scoping review aimed to identify and synthesize recent studies that used passive sensing techniques to extract patterns in daily routines and behaviours that could indicate loneliness or social isolation. Moreover, we performed a critical examination of these studies in various aspects like demographics, privacy considerations, and validation methodologies. As a result, this review provides us with a solid foundation for research by identifying the gaps in the existing literature related to passive sensing-based methods for loneliness detection, and based on those identified gaps, we have worked on subsequent studies.

2. The next contribution is about the development of personalized methods for loneliness detection to address the limitation presented in the section 1.5.1.2. To do so, we used clustering methods to identify similar behavioural subgroups in passive sensing data that have emerged over time and then developed personalized loneliness detection models for each subgroup to improve performance. A key feature of this work is the weekly analysis of the behavioural patterns of each identified group. Our results show that models tailored to student groups with similar behaviours are more effective at detecting loneliness than the one-size-fits-all approach, which highlights the importance of personalized detection systems.

3. Expanding upon the groundwork established in the last contribution, we introduced a personalized and dynamic approach for detecting loneliness to address the limitation presented in the section 1.5.1.3. The objective of this work is to have a dynamic loneliness detection method that adjusts to the user's evolving behaviours and continually improves its predictive capabilities with the integration of fresh data. We used incremental clustering to identify behavioural subgroups dynamically while, in parallel, applying incremental classification to detect loneliness in each subgroup. We also presented the evolution of weekly subgroups based on their behavioural patterns and loneliness levels. This helps in continuous monitoring and detecting behavioural changes that could indicate the emergence of loneliness.

4. The next contribution looked into the nuanced aspects of loneliness to address the limitation presented in the section 1.5.1.4. We used passive sensing data to explore social and emotional loneliness and its behavioural manifestations. This study investigates whether digital behavioural markers can differentiate between socially and emotionally lonely students, the predictive power of these markers in classifying socially and emotionally lonely students, and which markers are most important for these predictive models. This contribution improves the precision of loneliness detection and also provides a detailed understanding of its varied facets, which would lead to the development of interventions that are more closely aligned with the individual's specific experiences of loneliness.

5. The next contribution explores the complex relationship between loneliness and depression to address the limitation presented in the section 1.5.1.5. This study rigorously investigates how behavioural features captured by passive sensing correlate with and differ between loneliness and depression among students. It addresses important research questions to identify the common and unique behavioural characteristics of loneliness and depression through the lens of passive sensing. This helps us understand how these conditions influence each other and what specific behavioural features act as mediators within their interplay. Additionally, it evaluates the efficacy of using behavioural features alongside loneliness and depression scales to accurately classify individuals experiencing these conditions. This contribution provides perspectives on the behavioural indicators of loneliness and depression through passive sensing data.

1.7 Publications

There are 6 deliverables of this PhD. Four have been published, one has been accepted for publication, and one is currently being prepared for submission.

1. **Qirtas, Malik Muhammad**, Evi Zafeiridi, Dirk Pesch, and Eleanor Bantry White. "Loneliness and social isolation detection using passive sensing techniques: scoping review." JMIR mHealth and uHealth 10, no. 4 (2022): e34638.

2. **Qirtas, Malik Muhammad**, Dirk Pesch, Evi Zafeiridi, and Eleanor Bantry White. "Privacy-preserving loneliness detection: A federated learning approach." In 2022 IEEE International Conference on Digital Health

(ICDH), pp. 157-162. IEEE, 2022.

3. **Qirtas, Malik Muhammad**, Eleanor Bantry White, Evi Zafeiridi, and Dirk Pesch. "Personalising Loneliness Detection through Behavioural Grouping of Passive Sensing Data from College Students." IEEE Access (2023).

4. **Qirtas, Malik Muhammad**, Evi Zafeiridi, Eleanor Bantry White, and Dirk Pesch. "Evolving AI for Wellness: Dynamic and Personalized Loneliness Detection Using Passive Sensing." Accepted for Publication in Springer Lecture Notes in Computer Science as Proceedings of AISoLA Conference 2023.

5. **Qirtas, Malik Muhammad**, Evi Zafeiridi, Eleanor Bantry White, and Dirk Pesch. "The relationship between loneliness and depression among college students: Mining data derived from passive sensing." SAGE Digital Health 9 (2023).

6. **Qirtas, Malik Muhammad**, Evi Zafeiridi, Eleanor Bantry White, and Dirk Pesch. "Differentiating and Predicting Social and Emotional Loneliness in Students Using Passive Sensing: An Exploratory Study." In Preparation for Submission

1.8 Thesis Outline

In Chapter 2, we will present a comprehensive literature review on loneliness detection using passive sensing and will discuss different aspects of existing studies in this area. Chapter 3 presents the details of the datasets used, behavioural features extraction and overall methodology. Chapter 4 presents personalized detection through clustering and group-specific models. Chapter 5 presents a dynamic and personalized loneliness detection approach using incremental clustering and classification models. Chapter 6 explored social and emotional loneliness using passive sensing data. Chapter 7 investigated the relationship between loneliness and depression using passive sensing data. Chapter 8 concludes the thesis and presents limitations with future research directions.

Chapter 2

Loneliness and Social Isolation Detection through Passive Sensing Methods: A Comprehensive Overview

This chapter presents:

1. An in-depth look at the literature on the methods used to detect loneliness and social isolation. The aim is to bridge the gap between clinical methods and passive sensing methods for loneliness and social isolation detection.

2. A review of studies that used passive sensing technologies to detect loneliness and social isolation. We examined the type of populations that were studied, the different types of sensors that were used, and the efficacy and reliability of those methods.

3. Evaluation of this scoping review in connection with previous research, while providing a synthesis of the key findings and the implications for future research is also presented.

Passive sensing technologies provide an intuitive way of collecting data via personal devices such as smartphones and wearables. Scoping reviews, as compared to systematic reviews and other review-based techniques, are a relatively new method of searching and synthesizing the literature. They are useful for investigating broad areas across various study designs [MPP05, OWZS10], making them appropriate for examining emerging broader fields. There are four main reasons why researchers could conduct a scoping review in an unexplored or new area: to figure out the design and purpose of the study, to see if a systematic review is needed, to summarize the research results, and to find research gaps in the existing literature [AO05, LCO10]. This review investigated the field of loneliness and social isolation detection through passive sensing, which is a relatively new research area. Therefore there was a need to conduct a scoping review in this area.

2.1 Loneliness Detection Approaches: An Overview

2.1.1 Assessing Loneliness In Research And Clinical Settings

Traditionally, the detection of loneliness has been based on personal assessments in a clinical setting. This includes using the well-known University of California at Los Angeles (UCLA) loneliness scale and other validated measures by clinicians [Rus96, VHTC11, TK14, HD87]. Several other verified measures can be used to find out how lonely a person is [DJGK85, DJGVT99, GT06, GVT10]. These tools measure feelings of loneliness through an investigation of subjective experiences of isolation, support networks, and social interactions in the form of different questions. While these scales have played an important role in improving the understanding of loneliness, their reliance on self-reported data in a specific setting makes them vulnerable to inherent biases and limitations. Due to the subjective nature of these assessments and the fact that they are primarily verified through correlations with "at risk" populations, there is an urgent need for more objective, reliable, and thorough methods for detection.

2.1.2 Technology Based Detection through Passive Sensing

In recent years, there has been an increasing interest in using technology for more objective and unobtrusive monitoring of loneliness and social isolation detection. Passive sensing technologies are at the cutting edge of this advancement, and they made a shift in methodology from self-reported assessments to continuing, data-based assessments [TOO+19]. These technologies include a wide range of tools, such as ambient sensors in smart homes, smartphones, and wearable devices. For instance, smartphones can track movement patterns, social interactions (calls, messages), and app usage, while wearable devices can monitor physical activity, sleep patterns, and physiological metrics such as heart rate and skin conductance [CH18]. Ambient sensors in smart homes can detect daily routines, social interactions, and environmental factors [UKT+18]. These diverse data streams can be analyzed to find patterns and anomalies that correlate with loneliness and social isolation. The aim is to collect various types of users' data passively and then process that data to look for the signs or digital biomarkers in the behaviours of users that may indicate loneliness or social isolation. For example, reduced physical activity, irregular sleep patterns, and a decline in social interactions may serve as indicators of increasing loneliness.

1. **Smart-home-based Methods:** Smart-home technologies use ambient sensors to track everyday activities and well-being passively. This approach involves examining behavioural data, such as movement patterns, home appliance use, sleep schedules etc. to detect deviations from an individual's typical behaviour patterns that might suggest social isolation or loneliness indicators [HG08, WSS+13, SFJ+05]. For example, spending more time in certain sections of the home or abnormalities in sleep patterns may indicate social isolation or mental discomfort. Researchers have found that looking at these changes over time may be a more objective and long-term way to measure loneliness than using people to directly report their feelings [MCLP10, GM11, FGP08]. This method improves our understanding of loneliness patterns and allows for early intervention without requiring people to directly report or answer questions, thereby preserving privacy and minimizing biases.

2. **Mobile and Wearable Sensing Methods:** The increase in smartphone and wearable devices use has created new opportunities for monitoring health and well-being. These modern gadgets are equipped with a variety of sensors that may monitor physical activity, location, interactions with others (via call and SMS logs), and other data. By analyzing this sensor data, we can look for potential patterns of loneliness and social isolation like decreased mobility, less social engagement, and changes in communication habits. This method reduces the biases associated with self-reporting and gives insights into the constantly shifting dynamics of loneliness and social isolation in everyday life.

Table 2.1: Search Keywords and Result Statistics For Databases

Keyword Combinations	Databases					
	IEEE Xplore	Scopus	ACM	PubMed	WOS	EMBASE
Loneliness AND Detection	44	6487	1046	528	364	364
Loneliness AND Sensing	53	1128	1339	1153	1845	74
Loneliness AND Passive Sensing	3	94	64	21	38	4
Loneliness AND Monitoring	50	6443	1148	473	634	538
Social Isolation AND Detection	174	7329	432	2639	539	632
Social Isolation AND Sensing	78	834	946	2539	1473	74
Social Isolation AND Passive Sensing	1	127	21	43	29	4
Social Isolation AND Monitoring	135	7063	846	1485	1264	1164

2.2 Review Methodology

2.2.1 Search Strategy and Selection Criteria

A comprehensive literature search was conducted using the Preferred Reporting Items for Systematic reviews and Meta-Analyses (PRISMA) guidelines [LAT+09] for scoping reviews [TLZ+18]. The search included papers from January 2011 to January 2024. Given the interdisciplinary nature of our research, we included databases from both computer science and social science; IEEE Xplore, Scopus, ACM, PubMed, Web of Science, and EMBASE. The search focused on using passive sensing technologies to detect loneliness and social isolation. The search keywords and corresponding results for each database are included in the Table 2.1.

2.2.2 Data Extraction and Analysis

After an initial search, studies were selected based on specific inclusion and exclusion criteria. We selected studies that only focused on those using passive sensing technologies for detecting loneliness or social isolation. The data extraction from these studies involved examining the methodology, demographic characteristics, technology used, data collecting methods, and outcome indicators for loneliness or social isolation. This systematic method led to a comprehensive investigation of the effectiveness of passive sensing methods in different populations and settings, revealing important insight into the current research context in this developing field. The selection process is shown in Figure-2.1.

- **General description of the study:** Authors, year, country
- **Study design:** Population type and age, participants sample selection process, duration of the study, ground truth data collection methods, privacy handling.
- **Study tech insights:** Technology used for sensing, data collection streams, algorithm
- **Study outcome characteristics:** Indicators/identification markers, sensed outcomes.

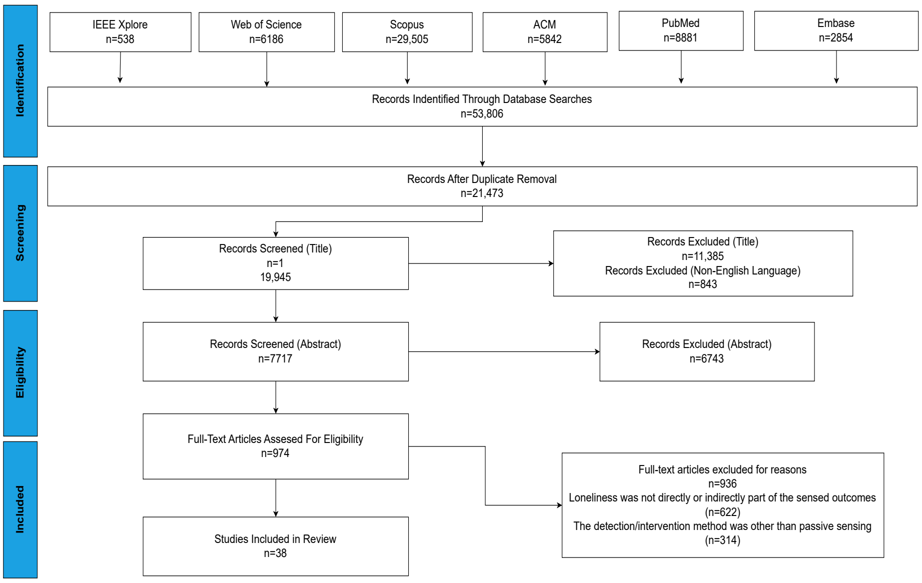


Figure 2.1: Flow diagram of literature search and selection process (adapted from Preferred Reporting Items for Systematic Reviews and Meta-Analyses [PRISMA]).

2.3 Results

2.3.1 Demographics Characteristics

2.3.1.1 Population Focus

Our review found a significant number of the studies focused on younger individuals mainly college students, which are 37% (14 out of 38) of the studies [WCC+14, DVC+19, JSW20, QGY18, PvHH+19, YLH+14, PA16, LSWL16, BZSW+15, WBC+21, EL23, YAJ+23, AA22, JAL+23]. Studies for older adults are 40% (15 out of 38) [GTT17, HPBGP17, WKL+14, PAK+13, MOEG17, ADR+16, VAAB+16, SMC+15, LAS+15, PTAK16, HTL17, LBD+21, RMP+23, JLY+23, SVA+22]. Only a smaller portion, 18% (7 out of 38), included participants from mixed age groups [SRRM+17, GNH+17, CRC+19, GLZ+16, FMG+20, WPB+21, SANV+22]. The number of participants in these studies ranged from as few as 5 to as many as 364, with study period extending from a week up to five years. While most studies employed a longitudinal design, collecting data over time to observe changes, one study [QGY18] used a cross-sectional design, examining participants only at a single point in time and two others did not specify their duration [LAS+15, CRC+19].

2.3.1.2 Age and Age Groups

40% (15 out of 38) of the studies explicitly mentioned the ages of the participants [QGY18, PvHH+19, WBC+21, WPB+21, PA16, WKL+14, GTT17, LAS+15, PTAK16, FMG+20, LBD+21, WBC+21, SANV+22, AA22, JAL+23]. The other studies referred to age groups in a general manner without mentioning specific ages. Four studies focused on individuals aged 18 to 28 years [WCC+14, DVC+19, JSW20, YLH+14, LSWL16, BZSW+15], whereas six studies focused on older adults [WKL+14, GTT17, LAS+15, PTAK16, FMG+20, LBD+21] and a few are mixed [SRRM+17, GNH+17, CRC+19, GLZ+16]. One study included a broad age range from 18 to 78 years [WPB+21], including both younger and older participants.

2.3.1.3 Gender Distribution

The gender distribution differed across the studies. Four Studies involving younger populations included both male and female participants [JSW20, YLH+14, FMG+20, WBC+21], 3 studies having more female participants [DVC+19, QGY18, WBC+21] and others 2 studies with male participants [PA16, BZSW+15]. Two studies solely consisted of females [PvHH+19, SANV+22]. Seven studies concerning older adults reported their gender, and all of these studies show a female predominance except one. The female ratios in these studies are 60% [WKL+14], 81% [ADR+16], 74% [GTT17], 54% [SMC+15], 89% [LAS+15], 44% [LBD+21] and 88% [PTAK16].

2.3.1.4 Educational Level

Participants' educational background was specified in 15 studies [DVC+19, PA16, YLH+14, LSWL16, WBC+21, BZSW+15, ADR+16, PTAK16, FMG+20, GTT17, LBD+21, WBC+21, SANV+22, EL23, AA22]. The younger participants mostly consisted of college students [DVC+19, PA16, YLH+14, LSWL16, WBC+21, BZSW+15, AA22, EL23, WBC+21], including first-year students and a combination of graduate and undergraduate students. Four Studies on older adults showed a diverse range of educational backgrounds [ADR+16, GTT17, PTAK16, LBD+21]. Two studies included participants who had completed their graduate studies [ADR+16, LBD+21], and the remaining two studies included participants with varying levels of education, with 46% completing a college education in one study [PTAK16] and 39% completing primary level education and 39% had no formal education in another study [GTT17].

2.3.1.5 Ethnic Background

The ethnic background of participants was reported in 10 studies, highlighting diversity, especially among younger populations [DVC+19, JSW20, YLH+14, LSWL16, BZSW+15, GTT17, LAS+15, LBD+21, EL23]. These younger groups often had an Asian cultural background, though studies also included participants from the United States [DVC+19, BZSW+15], China [LSWL16], and Singapore [YLH+14], representing a mix of ethnicities [JSW20, LSWL16]. Older adult studies reported a majority of Chinese participants in one study [GTT17] and a majority of White participants in another [LAS+15].

2.3.2 Sensing Streams

Smartphones and wearable fitness trackers are equipped with many sensors that help us track user movements and collect data on daily routines. The majority of the studies reviewed (58%, or 22 out of 38) used smartphones or wearable devices as the main method for passive data collection tool [WCC+14, DVC+19, JSW20, QGY18, PvHH+19, YLH+14, PA16, LSWL16, BZSW+15, WBC+21, MOEG17, SMC+15, SRRM+17, GNH+17, GLZ+16, FMG+20, WPB+21, GLZ14, WBC+21, RMP+23, YAJ+23, JAL+23]. The accelerometer and GPS are the most common sensors in these devices for passive data collection, with 72% (16 out of 22) of the studies using both of them [JSW20, QGY18, PvHH+19, GNH+17, CRC+19, BZSW+15, DVC+19, YLH+14, LSWL16, WBC+21, FMG+20, GLZ+16, SMC+15, RMP+23, YAJ+23, JAL+23]. Accelerometers are used because to their low power consumption and minimum privacy violation. GPS provides significant information on a user's travel routines, such as frequently visited places and the time spent at these places, which may help to extract social behaviours of users.

Bluetooth and microphones are the second most common sensors used in 45% (10/22) of the studies for passive data collection [DVC+19, PvHH+19, PvHH+19, LSWL16, WBC+21, GNH+17, BZSW+15, SMC+15, RMP+23, YAJ+23]. Bluetooth is used here as a proximity sensor which provides information about a user's sociability based on their physical closeness to others using Bluetooth-enabled devices. Microphones provide indirect indicators of sleep habits and social interactions. Additionally, a small number of studies (7%, or 3 out of 38) used Wi-Fi MAC addresses to collect social interaction data, and one study used a smartphone light sensor to determine screen lock/unlock state [GLZ+16], providing indirect information on device usage habits.

Phone usage data, along with hardware sensors, was shown to be an important source of information, serving as an indicator of social isolation or interaction. 50% (11 out of 22) of the studies used this approach to analyze SMS text messages and call logs to infer communication patterns [DVC+19, JSW20, YLH+14, MOEG17, PA16, LSWL16, SMC+15, GLZ+16, PTAK16, PvHH+19]. Some studies monitored how often smartphones were locked and unlocked [DVC+19, GLZ+16], while others analyzed app use patterns to infer mood and sociability. A study collected contact information, browsing history, and email data to get deeper insights into user behaviour [PA16].

34% (13 out of 38) of the studies shifted the focus from personal devices to ambient and physical sensors in houses [GTT17, HPBGP17, WKL+14, PAK+13, ADR+16, VAAB+16, LAS+15, HTL17, CRC+19, RMP+23, JLY+23, SVA+22, EL23]. This method focused on monitoring indoor mobility and behaviour patterns using a range of sensors, such as passive infrared motion sensors for tracking movement, pressure sensors for detecting presence on furniture, sound sensors for recording household activities and interactions, and door sensors for observing entry and exit patterns. This approach provides a different perspective to wearable and smartphone sensors, improving the comprehension of participants' lives and social interactions inside their homes.

2.3.3 Loneliness and Social Isolation Assessment: Methods and Tools

Multiple different methods of assessment were used to validate the data collected passively via sensors in the research of loneliness and social isolation. 72% of the studies (27 out of 38) relied on self-reports and questionnaires as their main methods for data collection. These tools provide a straightforward way to collect data on participants' health and well-being which could be either in a clinical environment or with using remote methods like through smartphone apps.

Self-reports and questionnaires were used to assess the physical and mental health of participants in 27% of the studies [JSW20, WBC+21, GTT17, MOEG17, GNH+17, GLZ+16, FMG+20, WPB+21]. These tools were crucial for assessing several aspects of social relationships, mental well-being, symptoms of anxiety or depression [JSW20, PvHH+19, LSWL16, BZSW+15, WKL+14, HTL17, LBD+21], and daily living activities [YLH+14, WBC+21, GTT17, SRRM+17, GNH+17, GLZ+16]. The selection of scales and questionnaires was customized to match the particular aspects of loneliness, depression, and anxiety. The UCLA loneliness scale [RPC80], the De Jong Gierveld Loneliness Scale [GVT10], and the ESTE-R Loneliness Scale [SMC+15] were often used for measuring loneliness. Similarly, for the evaluation of depression and anxiety symptoms used the Social Interaction Anxiety Scale, the Depression Anxiety Stress Scale, and the Positive Affect Negative Affect (PANA) Schedule [WCT88].

Some studies choose to use custom-designed questionnaires along with standard scales to meet their particular study requirements and target audiences. The ecological momentary assessment (EMA) is an example of a method that offers a customized way to record the nuances of participants' thoughts and behaviours as they occur. This approach enables researchers to gather data that is closely connected to their study's area and their requirements.

2.3.4 Efficacy and Reliability

2.3.4.1 Detection Methods

The studies reviewed used several indicators to identify changes in behaviour that indicate loneliness or social isolation. Of the indicators examined, the frequency of incoming and outgoing phone calls and SMS text messages is the most commonly studied, with 24% (9 out of 38) of the studies focusing on it [YLH+14, ADR+16, PA16, SMC+15, GLZ+16, PTAK16, JSW20, JAL+23, EL23].

Another often-studied indicator is how much time people spend away from their homes, with an equal number of studies investigating its influence on levels of loneliness [PAK+13, ADR+16, VAAB+16, GTT17, SMC+15, LAS+15, FMG+20]. Additional behavioural indicators examined include sleep cycles, voice activity and speech, physical movement levels, mobility and pace of walking, time spent in various areas of the home, types of applications used, and interactions with others identified through Bluetooth proximity detection [WKL+14, DVC+19, GTT17, BZSW+15].

2.3.4.2 Phone Calls or SMS Text Messages

Nine studies investigated the relationship between the frequency of incoming and outgoing phone calls and the feeling of loneliness [GLZ+16, PTAK16, SMC+15, LSWL16, WPB+21, MOEG17, PA16, ADR+16, EL23]. One significant finding was that those who received fewer incoming calls felt lonelier than those who received more calls [GLZ+16]. One study with older adults showed that the loneliest person got just 40% of the phone calls compared to the least lonely person [PTAK16], indicating a significant difference in social interaction. The influence of communication between family and friend circles was also examined closely. A study found that in intimate relationships, both receiving and making calls to family members are important factors in determining feelings of loneliness inside the family and between spouses [SMC+15]. In comparison to social isolation, the frequency of phone calls from acquaintances, as opposed to friends, was shown to be a more reliable indicator. This detailed comprehension helps in creating prediction algorithms to identify loneliness. Upon further examination, it was found that those with increased symptoms of loneliness received fewer phone calls on average [PTAK16]. An analysis demonstrated a slight inverse relationship between the duration of phone conversations and loneliness, indicating that lengthier calls may be linked to reduced feelings of loneliness.

Additionally, the length of phone calls was shown to have an inverse relationship with social well-being, which could mean that long chats might have a beneficial impact on a person's social health. Two other studies also confirmed this result by identifying a negative relationship between call length and loneliness [MOEG17, PA16]. The frequency of calls was significantly linked to the overall phone use time [ADR+16]. A study on students found that those who received more phone calls, particularly on weekdays, reported reduced feelings of loneliness, an easier adjustment to college life, and a better sense of community among their peers [YLH+14]. Research indicates that those with increased levels of social anxiety or loneliness often get a lower number of SMS text messages [GLZ+16]. The study indicated a little negative association between receiving SMS messages at night and loneliness. However, there was no significant connection between the frequency of messaging and feelings of loneliness or social well-being [WPB+21]. A study found a strong connection between the number of messages received and scores on the UCLA loneliness scale which explains the impact of digital communication on how loneliness is perceived and felt [PA16].

2.3.4.3 Web Social Activity and Communication Apps

Five studies investigated the correlation between online social activity and loneliness [LSWL16, WPB+21, GLZ+16, YAJ+23, AA22]. A study conducted with a younger population found a positive correlation between loneliness and the frequency of app usage [LSWL16], suggesting that younger people who spend more time on social applications tend to feel more lonely. In comparison, a study on older people found that those who increased their social media use saw a significant decrease in symptoms of loneliness [WPB+21]. Researchers discovered that the social media networks of older people resulted in less loneliness and higher levels of overall well-being when compared to those of younger adults. Another study discovered that smartphone activities connected to web-based social interactions, such as using social or communication apps often, did not show any connection to feelings of social anxiety or loneliness [GLZ+16].

2.3.4.4 Number of Computer Sessions or Hours on the Computer

Furthermore, 2 studies examined computer use based on the frequency of sessions and total daily screen time [ADR+16, LBD+21]. An increasing number of computer sessions was found to be linked to increased feelings of loneliness. Previous research examining the correlation between computer usage and loneliness have reported conflicting results [MMS03]. While some argue that using computers might help relieve loneliness, others suggest that more computer use, especially among young individuals is linked with increased feelings of loneliness [FK07, Cap06, CC08, MM99]. A study found no significant correlation between phone usage, computer use, and loneliness [ADR+16]. The number of phone calls was significantly linked to the time spent on the phone, whereas the number of computer sessions had a strong connection to the time spent on the computer [ADR+16].

2.3.4.5 Bluetooth Sensing

Five studies investigated the correlation between Bluetooth proximity sensing features related to social interactions and feelings of loneliness [WBC+21, WCC+14, LSWL16, DVC+19, WBC+21]. A study found that individuals often report experiencing loneliness when many Bluetooth devices are detected nearby. The tendency might indicate instances when an individual experiences solitude in crowded public places with a significant number of Bluetooth devices present [WBC+21]. Another study found a significant relationship between Bluetooth device interactions and the Patient Health Questionnaire-9 depression scale [WCC+14]. A study found that there is no correlation between Bluetooth proximity and loneliness [LSWL16]. A study used Bluetooth along with other sensor data including GPS, Wi-Fi, SMS text messages to improve the accuracy of predicting loneliness levels in individuals [DVC+19].

2.3.4.6 Daily Phone Use or Time on Mobile Phone

Three studies examined the correlation between overall mobile phone use and loneliness [PTAK16, DVC+19, LSWL16]. A study found that less phone use on certain weekends and in the morning was linked to higher levels of loneliness in younger individuals [DVC+19]. Another study found that students who are at risk of loneliness tend to spend more time on their smartphones daily. Loneliness was shown to be positively linked to app use, regardless of the time of day [LSWL16].

2.3.4.7 Demographic Characteristics

Out of the 38 studies, 6 studies examined differences between genders in loneliness levels [PTAK16, GLZ+16, GLZ14, WPB+21, RMP+23, SVA+22]. Petersen et al [PTAK16] found that gender significantly influenced the daily call frequency, with women making or receiving twice as many calls as males. A study investigated the gender distribution between high and low loneliness groups. The chi-square test showed that the gender-based difference was not statistically significant [GLZ+16]. A study found that women with higher levels of well-being use reading applications and browsers more often than males [GLZ14]. Four studies investigated passive sensing and participant behaviour in relation to age. A study investigated the ages of individuals in high and low loneliness groups; the chi-square test showed that there was no statistically significant age difference [GLZ+16]. Another study found that age did not significantly affect the daily number of calls [PTAK16]. Wetzel et al found a significant correlation between participants' age and time spent on social media [WPB+21], showing that the link between social media usage and loneliness differs depending on age. Older individuals who spend more time on social media are likely to feel much less lonely. While younger people who spend more time on social media were linked to higher levels of reported loneliness [WPB+21]. Another study discovered that the level of activity and the frequency of messages and calls increased in relation to the age of individuals born prior to the year 2000 [SRRM+17]. Cognitive ability was considered a demographic variable in 2 studies. A study found that those with a chronic disease have a higher level of loneliness and more severe social well-being compared to those without a chronic illness [WPB+21]. Another study found a correlation between improved cognitive abilities and increased frequency of daily phone calls [PTAK16]. This study investigated whether an individual's pain level may predict phone calls but found no significant correlation between pain level and call frequency has been found [PTAK16]. Two studies investigated the correlation between an individual's personality characteristics and activity patterns inferred from smartphone data collected passively [PA16, SRRM+17]. A study found a strong connection between emotional stability and extroverted personality traits and several smartphone-based features. Agreeableness, conscientiousness, and intellect personality traits are moderately linked to various smartphone activities, including messaging, browsing, calls, and contacts [PA16]. Another study found that thankfulness is linked to participants' communication habits, with the most thankful individuals mostly using their phones to communicate.

2.3.4.8 Time Outside of Home and Time Spent at Home

Seven studies evaluated the impact of time spent outside compared to at home on an individual's loneliness levels [LAS+15, FMG+20, GTT17, ADR+16, HTL17, PAK+13, VAAB+16]. A study with older people found that more time spent outside the house was linked to reduced feelings of loneliness [LAS+15]. Other studies including participants of various age groups found that those who increased their time spent outside noticed a decrease in feelings of loneliness [FMG+20]. Previous research studies have examined outdoor activities among older individuals by assessing their average daily time spent outside and the frequency of outdoor visits. Spending more time outside linked to reduced loneliness and increased social networking scores. Furthermore, two more studies found a significant correlation between the amount of time spent outside daily and feelings of loneliness [ADR+16, HTL17]. Another study found that those who spend a significant amount of time at home have higher feelings of loneliness. The physical activity score correlates positively with the average time spent outside the home [PAK+13].

2.3.4.9 Daily Activity and Movement

Five studies investigated how everyday activities and movements, as well as their duration, influence levels of loneliness [BZSW+15, FMG+20, LSWL16, DVC+19, GTT17]. A study found that individuals who engage in higher levels of physical activity tend to report lower levels of loneliness [BZSW+15]. In another study, a higher level of dispositional loneliness was linked to a significantly shorter average activity duration [FMG+20]. Feelings of detachment from others and painful sentiments of isolation characterize dispositional loneliness [RPP+20]. Loneliness was shown to correlate positively with the frequency of visits to significant places [FMG+20]. Loneliness is inversely associated with activity duration [FMG+20]. A study with students found a significant negative relationship between loneliness and indoor mobility, such as the time spent moving inside throughout the day. Studies show that sedentary students are at a higher risk of feeling lonely. Another study found that those who participate in activities for longer periods experience less loneliness, regardless of the time of day or night [LSWL16]. Another study found the same trend: increased overall physical activity and number of steps throughout the day and night lead to reduced feelings of loneliness [DVC+19].

2.3.4.10 Conversation Activity

Three studies investigated how conversation and duration of conversations impact levels of loneliness [BZSW+15, FMG+20, WCC+14]. Out of these, two studies reported no significant relationship between conversation duration and loneliness [BZSW+15, FMG+20]. While the other study found that college students who had fewer social interactions are more prone to loneliness [WCC+14].

2.3.4.11 Sleep

Three studies examined the correlation between the duration of sleep and feelings of loneliness, depression and stress [BZSW+15, WCC+14, GTT17]. A study found that there was no correlation between sleep duration and loneliness, but there was a negative relationship between sleep duration and daily stress [BZSW+15]. This indicate that those who sleep enough tend to have lower levels of stress. A study with students found that those who get less sleep are more prone to depression [WCC+14]. Please note that this study didn't explicitly analyze the relationship between sleep and loneliness. A study on older people revealed that those who noticed themselves as socially isolated had more midday naps [GTT17].

2.3.4.12 Privacy Issues and Ethical Concerns

62% (23/38) of the studies did not discuss privacy or ethical issues and did not specifically mention privacy concerns [SRRM+17, HTL17, HPBGP17, DVC+19, MOEG17, ADR+16, GNH+17, PA16, LSWL16, VAAB+16, GTT17, CRC+19, SMC+15, GLZ+16, BZSW+15, LAS+15, WBC+21, PTAK16, FMG+20, WBC+21, RMP+23]. Five studies included details about ethical clearance from the appropriate committees and participant consent [QGY18, PvHH+19, WKL+14, WPB+21, JSW20]. Two studies reported that they stored user data on a secure data server and used anonymized user IDs to protect users' privacy [WCC+14, BCC+21]. Furthermore, a study collected statistical data on smartphone use only to protect the anonymity of participants [YLH+14]. Another study used a video camera installed in a smart home to monitor users' activities. They recorded video clips for only 5 seconds each time activity was detected at the entrance [PAK+13].

2.4 Discussion

2.4.1 Comparison With Previous Work

Our scoping review contributes a new perspective to the existing literature on passive sensing by being the first one to focus on detecting loneliness and social isolation using passive sensing methods. Previous reviews have thoroughly discussed the use of passive sensing in monitoring different mental and physical health conditions. Our study addresses a gap in the literature by specifically examining loneliness and social isolation. This emphasis highlights the current challenges and possible future directions in the area, especially the limitations of current passive sensing techniques in reliably identifying these complex conditions. Existing studies on loneliness detection using passive sensing still have significant gaps. This thesis aims to contribute to filling some of these gaps.

2.4.2 Target Group Selection and Technology Choice

Selecting the target group is an important aspect for developing systems that detect loneliness or social isolation. Our review found that the majority of studies primarily focused on younger age groups. This inclination may come from the practical convenience of having a younger group, especially from educational institutions like colleges and universities, as opposed to the more complicated challenge of including older people, which is often spread out throughout different community scenarios, which makes them harder to reach via traditional channels. Existing studies support the observation that isolation tends to peak during the transitional phase of late adolescence and again in the post-retirement period, which highlights a distinct age-related pattern in social isolation and loneliness. Middle-aged individuals, particularly those who are jobless, are recognized as another demographic with a higher vulnerability to loneliness. The varying vulnerability at different life phases highlights the need for a precise selection of detection technologies customized for each age group. Another important observation from this review is that the methods used for loneliness assessment are different for different age groups. Younger people are primarily studied through the lens of smartphone usage, which may be based on assumptions about their higher engagement with digital devices and platforms. In contrast, studies focusing on older adults more frequently used ambient sensors for passive data collection; this could be possibly due to the less use of smart devices by this population group, and usually, older people are more critical about privacy issues. Many factors could contribute to the choice of using ambient sensors for older people. Firstly, many older people aren't comfortable with technology and find the interfaces on smart devices too complicated to use. Additionally, individuals suffering from conditions like dementia could face other practical challenges related to using such modern devices, such as remembering to carry, wear, or charge these devices etc. The difference in technology acceptance and preference for detecting loneliness between younger and older people highlights the need of creating systems that are suitable for different age groups and easy to use. Developers may improve the effectiveness of loneliness detection systems by recognizing and overcoming particular challenges encountered by older persons, therefore customizing solutions to accommodate the different needs of individuals based on the age group.

2.4.3 Privacy Issues

As we have seen, privacy concerns were one of the important aspects among participants. While a few studies mentioned taking steps to ensure privacy, such as obtaining participant consent for data collection and securing ethical clearance but there is still a significant portion of studies that did not explicitly discuss how privacy concerns were addressed. This oversight raises important questions about the overall acceptability of these detection methods to both participants and the general public, as well as potential risks associated with data and privacy breaches.

2.4.4 The Need for Personalized and Dynamic Loneliness Detection

In most of the studies reviewed, researchers relied on generic and static machine learning models that use all the available data to train a model for loneliness detection. This approach does not effectively account for the nuanced and diverse ways in which loneliness manifests across individuals. This also lacks the adaptability to capture the evolving nature of individual behaviours and the fluctuating nature of loneliness itself. By relying on a single, generalized model trained on the entire dataset, it may overlook subtle but crucial differences in behavioural patterns between individuals or subgroups. For example, it may not account for variations in phone usage patterns across different personality types or social contexts which could lead to misclassification of individuals with unique behaviours. Furthermore, generic models may struggle to capture the dynamic nature of loneliness, which can fluctuate over time and be influenced by various contextual factors. This is particularly relevant in the context of college students, where daily routines, social interactions, and individual preferences vary. There is a need for methods to recognize the heterogeneity within the population and then develop personalized models for loneliness detection that can effectively capture the unique ways in which loneliness is expressed, leading to improved detection accuracy and, ultimately, more targeted and effective interventions.

2.4.5 Distinguishing Social and Emotional Loneliness

There are a number of studies in the psychology literature that have examined types of loneliness, social and emotional loneliness [MVVdNG17, QM02, Zam08, OKDE+22], but there is not any study that investigates these types through passive sensing. While both types collectively contribute to the overall experience of loneliness, they have distinct facets with potentially unique behavioural manifestations [QM02]. Investigating these nuanced aspects of loneliness is important for finding their unique behavioural manifestations and differences. This can lead to improved precision of loneliness detection and more personalized interventions that address the specific needs of individuals experiencing a specific form of loneliness.

2.4.6 The Loneliness-Depression Relationship

Loneliness and depression are interconnected conditions and often co-occur while influencing each other in complex ways [SK13]. Studies have shown that loneliness can lead to depression and vice versa, creating a vicious cycle that exacerbates both conditions [CHT10, LWGB15]. All the reviewed studies in this scoping review have exclusively focused on loneliness detection using passive sensing, while there exists a separate body of literature that independently explores the use of passive sensing for depression detection [NGT+20, JC20, DALW+22, MRS+22]. No studies, to our knowledge, have explicitly investigated the relationship between loneliness and depression through the lens of passive sensing data. This highlights the need for studies that use passive sensing to explore the bidirectional relationship between loneliness and depression to provide a more comprehensive understanding of their interplay.

2.4.7 Energy Consumption and Validation Challenges

Energy consumption emerged as another critical concern. Continuous data collection from various sensors can drain a smartphone's battery quickly, which is inconvenient for users and can interrupt data collection. However, only a few studies mentioned this issue, either by planning to tackle it in the future or by implementing strategies like lowering the sensors' sampling rate when the battery is low or selectively using sensors that consume less power.

The comprehensive validation of loneliness detection systems before their public deployment is important to ensure that these systems are not only accepted by targeted users but also trusted for their accuracy, reliability, and sensitivity to privacy concerns. This validation process is critical as it provides key information about the systems' functionality and efficacy.

2.5 Conclusion

This chapter has highlighted several key gaps and challenges in the current research on loneliness and social isolation detection using passive sensing. The rest of this thesis aims to address some of the gaps discussed above, including personalized and dynamic loneliness detection, investigating social and emotional loneliness using passive sensing, and exploring the loneliness and depression relationships through the lens of passive sensing.

Chapter 3

Methodology

3.1 Smartphone Sensors for Passive Sensing

In today's world, smartphones have become essential tools for everyday activities. Beyond their primary function as a communication device, smartphones are equipped with several sensors. These sensors are a powerful tool for passive data collection [EGH+14]. Passive sensing has a significant role in psychological and behavioural research. It enables data collection in a realistic setting, which can be used to obtain data about individuals' real-world behaviours and feelings [NA18]. Each sensor on a smartphone can capture different types of data, which helps in getting unique information about user behaviour and their environment. The following are the key sensors from smartphones that have been used in this thesis.

1. Location (GPS): The GPS sensor provides precise geographical coordinates. Location data can be important in understanding an individual's mobility patterns and social tendencies. Frequent time spent in isolated places or reduced mobility can indicate social withdrawal which is a common symptom of loneliness [WN12].

2. Bluetooth: Bluetooth radio can detect the presence of other nearby Bluetooth-enabled devices. Bluetooth can help to provide information about social interactions and networks by capturing data on nearby devices. It can indicate the frequency of an individual's proximity to others which can provide information about their social engagement levels. A decrease in detected devices might indicate reduced social engagement which is again an important indicator of loneliness.

3. WIFI: Similar to Bluetooth, WiFi radio detects WiFi networks and their strength. WiFi connectivity patterns can be used to infer an individual's location within a certain area, like home or work, and their movement patterns within these environments.
4. Calls/SMS Logs: These logs provide data on communication patterns, including call duration, frequency, and messaging activities. Communication logs are an important source of information for finding social interaction patterns. Decreased frequency or changes in communication habits can indicate changes in social behaviour, which could be linked to feelings of loneliness or isolation.
5. Phone Usage: This involves tracking the usage patterns of smartphones. Screen time data can provide information about daily routines, sleep patterns, and potential over-reliance on digital devices for social interaction.
6. Conversation: This uses a microphone sensor on a smartphone to record the start and end timestamps of conversation activity. Again it's an important indicator for social communications.

3.2 Fitness Bands and Wearable Technology

The popularity of fitness bands and wearable technology has revolutionized the field of passive sensing in monitoring physical and mental health related indicators [SQK21]. These devices are widely accessible and can provide continuous, real-time data, which makes them important tool in behavioural research. Fitness bands are equipped with several sensors that can track multiple aspects of physical health which can provide important information about the user's lifestyle, habits, and even mental well-being [CH18]. Below are the fitness band sensors that have been used in this thesis.

1. Physical Activity: Fitness bands typically include accelerometers and gyroscopes to measure movement and physical activity like step count. Low daily physical activity can indicate some physical or mental health condition [PDA05].
2. Sleep: Wearables use a combination of movement detection and heart rate monitoring to track sleep patterns, including duration and quality. Sleep is a critical component of overall health and well-being. Monitoring sleep patterns can give us data about disturbances or irregularities that are often

associated with various psychological conditions, including stress, anxiety, and loneliness [AB13]. Consistent sleep disturbances or significant variations in sleep patterns could be indicative of underlying mental health issues.

3.3 Datasets

This thesis mainly used two existing datasets collected through smartphones and fitness band sensors. These datasets were selected based on their availability and their relevance to addressing the key research questions. The StudentLife dataset was chosen for the initial studies because it was the only available dataset at the time that provided the necessary passive sensing and survey data related to mental health. It also has been used in existing literature as reviewed in Chapter 2. Later, the GLOBEM dataset became available, which provided more recent data and a larger participant population, making it well-suited for investigating subsequent research questions.

3.3.1 StudentLife Dataset

The StudentLife dataset is a collection of data that was passively collected from undergraduate and graduate students at Dartmouth College in 2013 [WCC+14]. It provides data about students' daily behaviours and mental health using different smartphone sensors, multiple clinical scales, and different ecological momentary assessment (EMA) responses. The dataset collection study received ethical approval from Dartmouth College's Institutional Review Board.

We used the first dataset from the StudentLife study, which had 48 students. The dataset was collected over the period of 10 weeks during the spring semester. From the comprehensive dataset, we selected a subset of sensors based on their relevance to loneliness detection as identified in existing literature [DVC+19]. These sensors include activity, location, call and SMS logs, Bluetooth, WIFI, conversation and phone usage data.

For loneliness measurement, this dataset has used the University of California Los Angeles (UCLA) scale [RPC80]. It is a 20-item questionnaire that provides the ground truth for loneliness levels by generating a continuous score that reflects participants' self-reported loneliness. While passive sensing data provides objective information about behavioural patterns, clinical scales like the UCLA Loneliness Scale provide the real ground truth data required to label the behavioural features to train and validate machine learning models as we discussed in detail in Chapter 1. The UCLA scale includes both positive and negative items, with the total scores ranging from 20 to 80. Positive items ask about feelings of connection or belonging, while negative items focus on feelings of isolation or disconnection. For example, a positive item is "I feel like I am part of a group of friends," while a negative item is "I feel left out". Participants rate how often they experience each feeling which leads to a cumulative overall score of loneliness experience. We used the threshold of 43 as the cutoff for loneliness, based on prior studies [DVC+19, CP08]. Participants scoring above this threshold are considered to experience significant loneliness. Within the dataset, loneliness scores varied widely among participants, ranging from as low as 25 to as high as 64 with 12 out of 41 students falling into high loneliness category. This scale was administered at two stages: pre and post-study. Table 3.1 presents the demographic and other related information for StudentLife dataset.

Table 3.1: Study Information and Participant Demographics for StudentLife Dataset

Category	**Data**
Participants	Total: 48
	Undergraduate: 30
	Graduate: 18
Class Year	Seniors: 8, Juniors: 14, Sophomores: 6, Freshmen: 2
	Ph.D.: 3, Masters (2nd year): 1, Masters (1st year): 13
Gender	Female: 10, Male: 38
Ethnicity	Caucasian: 23, Asian: 23, African-American: 2
Ground Truth	Pre-study: 20-item UCLA scale, Post-study: 20-item UCLA scale
Sensor Data	Activity, Location, Call , SMS Logs
	Phone Usage, Conversation, Bluetooth, WIFI

3.3.2 GLOBEM Dataset

The GLOBEM dataset is a collection of passively collected data from undergraduate students over a period of 10 weeks at the University of Washington during the spring semester of 2019 [XZS+22]. This 10 week period from March to June was specifically chosen to assess any potential semester effects on the students behaviours. A total of 218 undergraduate students, with a majority being female, participated in this study. These students were recruited through email and social media invitations, and detailed demographic information about the participants is provided in Table 3.2.

Table 3.2: GLOBEM Dataset Participant Demographics. Gender acronyms: F: Female, M: Male, NB: Non-binary. Racial acronyms: A: Asian, B: Black or African American, H: Hispanic, N: American Indian/Alaska Native, PI: Pacific Islander, W: White, NA: Did not report. The & symbol denotes participants who identified with multiple races.

Category	**Data**
Participants	Total: 218
	Gender: F 111, M 107
	Ethnicity: A 102, B 6, H 10, N 2, PI 1, W 70,
	A&B 1, A&W 16, H&W 2, B&W 2,
	A&H&W 1, B&H&W 1, H&N&W 1, NA 3
Ground Truth	Pre-study 10-items UCLA scale
	Post-study 10-items UCLA scale
Sensor Data	Bluetooth, WIFI, Call Logs
	Location, Phone Usage, Physical Activity, Sleep

The study used a smartphone app developed using AWARE framework for data collection that was designed to run continuously in the background and collect data without needing any engagement from the users. The data collection frequency is different for different sensors and details can be seen on the AWARE official site here [Fra24]. A Fitbit device was used to collect data on the sleep patterns and physical activity of the participants. The dataset includes nearby logs of Bluetooth and WIFI devices encountered, location, phone usage, calls, physical activity, and sleep data.

The University of Washington's Institutional Review Board (IRB) granted ethical approval for this data collection. All participants in the study provided informed consent. Any information which could directly identify participants was restricted to the core data team and data has been removed for any participant who chose to withdraw from the study.

Loneliness levels among the participants were assessed using the revised 10-item UCLA loneliness scale [KCMG88]. This scale required participants to rate their feelings of loneliness on a scale from 1 ('never') to 4 ('always'), with reverse scoring applied to five of the items. The scores from these items were then aggregated which gives a total score range between 10 and 40, where higher scores indicate greater feelings of loneliness. This scale was used pre- and post-study. Since the revised 10-item UCLA scale is less commonly applied in existing studies on loneliness detection, there is no widely established standard cutoff score in the literature. After reviewing the distribution of scores and considering the characteristics of the scale, we selected a cutoff score of 20. This threshold was chosen to align with the mid-point of the scale, where scores of 20 and above would indicate moderate to high loneliness based on the participants' self-reported experiences. 113 out of 218 students fall into the lonely category.

3.4 Behavioural Features Extraction

The main goal of feature extraction in the context of loneliness detection using passive sensing data is to transform raw sensor data into a meaningful set of features that could be used for further analysis and for training models. We extracted behavioural features using the Reproducible Analysis Pipeline for Data Streams (RAPIDS) [VLA+20]. It is a framework that has been developed specifically for extracting digital behavioural features through passively collected datasets. A detailed overview of these features is provided in RAPIDS documentation [VLA+20]. For a summarized form of the extracted features, please refer to Table 3.3. A sample of the feature matrix schema is provided in Fig.3.1. Each row in the matrix represents a daily data sample for a participant, identified by a unique 'pid' (Participant ID) and date. The columns represent various features extracted from different sensors. Since people engage in different activities throughout the day, the features were extracted for different time epochs (morning, day, evening, night) alongside the day-level (24-hour) features. Sleep related features were extracted over the entire sleep period rather than divided into epochs.

pid	date	f_wifi:phone_wifi_connected_rapids_uniquedevices_dis:morning	…	f_call:phone_calls_rapids_incoming_meanduration:allday	…	f_slp:fitbit_sleep_summary_rapids_sumdurationasleepmain	…	f_steps:fitbit_steps_summary_rapids_maxsumsteps:allday	…
1	21/03/2019	3	…	15	…	1369	…	24530	…
2	22/03/2019	1	…	23	…	1123	…	845	…
…	…	…	…	…	…	…	…	…	…
20	21/03/2019	6	…	130	…	954	…	1139	…
…	…	…	…	…	…	…	…	…	…

Figure 3.1: The sample schema of the feature matrix (each column is a feature and each row is a daily sample per participant)

Table 3.3: Summary of Sensor-Derived Features; StudentLife (SL), GLOBEM (GB)

Category	**Derived Sensed-behaviours**
Bluetooth (SL, GB)	Number of scans, unique devices, mean and standard deviation of scans per device, most and least frequent device scans within segments, across segments, and across the entire dataset.
Wi-Fi (SL, GB)	Number of scans, unique devices, and scans of the most connected device measured over different time segments
Location (SL, GB)	Total distance traveled, radius of gyration, maximum diameter, maximum distance from home, significant locations visited, distance lengths and durations (mean and standard deviation), Shannon entropy, circadian routine, location variance
Phone Usage (SL, GB)	Total, longest, shortest, average, and standard deviation of unlock duration, number of unlock episodes, time until first unlock
Call (SL, GB)	Number of calls, distinct contacts, mean, sum, minimum, maximum, standard deviation, mode, and entropy of call durations, time of first and last call, calls with the most frequent contact.
Physical Activity (SL, GB)	Maximum, minimum, average, median, and standard deviation of daily steps, total steps, number, and duration (total, max, min, average, standard deviation) of sedentary and active bouts.
SMS (SL)	Total number of messages, number of unique contacts, time of first and last message, number of messages from the most frequent contact.
Conversation (SL)	Total, maximum, minimum, and average duration of conversations; standard deviation of conversation duration; time of first and last conversation
Sleep (GB)	Number, total, longest, shortest, average, median, and standard deviation of awake/asleep episodes (main, nap, all), first and last wake/bedtimes, sleep efficiency, total and average duration to fall asleep, asleep, awake, in bed, and after wakeup for each sleep type.

3.4.1 Location Features

Location features extracted from smartphones provide information about participants' movement patterns and social behaviours, which are important indicators in loneliness detection. These features are derived from GPS data and processed to understand how participants interact with their environments. Below are some of the key features.

- **Total Distance and Speed:** This computes the total distance travelled by a participant within a specific epoch. Additionally, it calculates the average speed during periods labeled as 'Moving', as well as the variance in speed, to understand the variability in their movement.
- **Significant Places and Transitions:** The number of significant (unique) places visited gives an insight into the participants' social habits. It identifies these locations using clustering algorithms and tracks the number of transitions between these places to capture social mobility.
- **Predictability About Locations:** These are features about location entropy. This uses the Shannon Entropy (measures uncertainty or randomness in data) to measure the predictability of participants' location patterns. Higher location entropy indicates more varied and less predictable movements, while lower entropy means more consistent routine and predictable location behaviours.

3.4.2 Phone Usage Features

Phone usage features extracted from passive sensing data provide data on how individuals interact with their smartphones. These features are critical for loneliness detection as they can provide information about social connectivity, mood fluctuations, and overall mental well-being. For example, frequent phone usage might indicate a search for social interaction, while irregular usage patterns could reflect emotional distress or social withdrawal [LTC14, GLZ+16]. By analyzing these features, we can better understand the relationship between phone usage behaviours and feelings of loneliness. Below are some key features.

- **Total Duration of Use:** This sums up the total minutes spent during all phone unlock episodes in an epoch. A 'phone unlock episode' means any instance when a user activates their phone screen from a locked state. This total duration helps us understand the overall time a user engages with their phone.
- **Unlock Episode Duration:** This looks at the longest and shortest duration of phone use in any single unlock episode. This also calculates the average duration across all episodes. These metrics provide a sense of how users interact with their phones; whether they have prolonged periods of use or just brief checks.
- **Variability in Usage:** The standard deviation of the duration of all un-

lock episodes indicates the consistency of phone use. High variability may indicate irregular usage patterns which could be associated with changes in mood or social activity.

- **Frequency of Use:** By counting the number of unlock episodes, this measures how frequently the user interacts with their phone. Frequent unlocking could mean a search for social connections or a response to received notifications.

3.4.3 Conversation Features

Conversation features are extracted through the passive data collected through the microphone of an individual's smartphone. It provides the start and end timestamps of each verbal communication activity. These features can provide important information about social engagement and verbal communication patterns. Below are some key features.

- **Conversation Duration and Frequency:** This calculates the total duration of conversations, the longest, shortest and average durations.
- **Conversation Timing:** This calculates the times of the first and last conversations in a day that can indicate a person's daily communication rhythms.

3.4.4 Physical Activity/Steps Count Features

Physical activity and step count features provide information about an individual's physical activity levels, which is an important indicator of overall well-being. These features are particularly relevant for loneliness detection as variations in physical activity can reflect changes in mood, social interactions, and daily routines [Mor13]. Below are some key features.

- **Sum/Avg Duration:** Measures the total/Avg time spent in each activity state (active, sedentary) within a time segment.
- **Range of Steps:** It extracts the highest and lowest number of steps taken by an individual within a time segment.
- **Average Daily Steps:** It calculates the average number of steps taken each day during a time segment which provides a baseline measure of an individual's general activity level.

- **Median Daily Steps:** It calculates the median step count, which can be less affected by outliers in the data and may offer a more representative value of typical daily activity.
- **Variability in Daily Steps:** It measures the standard deviation of the daily step count throughout a time segment to understand the consistency of physical activity. A higher standard deviation can indicate significant variations in daily activity which could be linked to changes in routine or social patterns.

3.4.5 Call Log Features

Call log features extracted from smartphones provide important information about an individual's social interactions and communication patterns. These features are important for loneliness detection as they can tell about the frequency, diversity, and nature of social connections. Frequent and diverse communication may indicate strong social ties, while irregular or limited call patterns could indicate social isolation or emotional distress [TR21]. A low frequency of calls and limited contact diversity can indicate a smaller or less active social network which is an important indicator of loneliness [SHC+21]. Below are some key features.

- **Call Frequency and Contact Diversity:** Metrics include the total number of incoming and outgoing calls in an epoch and the count of unique contacts.
- **Call Duration:** It calculates the average, total, shortest, longest, and the range of call durations. The mode and entropy are also computed to find the regularity and diversity in call durations.
- **Missed Call Patterns:** Different features related to missed calls like looking at how often they occur, the variety of contacts that calls are missed from, and the timing of these missed calls.

3.4.6 Messages/SMS Log Features

Similar to call log features, messages are also an important element in finding an individual's social engagement, communication patterns, and potential signs of social withdrawal. However, it is important to note that the use of traditional SMS messaging has reduced significantly in recent years due to the more common usage of messaging apps like WhatsApp. But still in the context of the

StudentLife dataset (collected in 2014), SMS data remains a relevant indicator of social interaction. Below are some of the key features.

- **Message frequency and contact diversity:** This Includes the total number of text messages sent and received, as well as the count of unique contacts involved in the messaging per each time segment. These figures are important indicators of the level of social engagement of an individual's social network.
- **Message timing:** By recording when the first and last messages are sent each day and during specific periods, can provide a person's daily communication habits.

3.4.7 Connectivity features: Bluetooth and WiFi

Bluetooth and WiFi connectivity features can provide data about an individual's social environment and potential interactions. These features include the frequency of device encounters, the diversity of connections, and the persistence of interactions with specific devices or networks. Analyzing these features can help researchers identify patterns that may correlate with loneliness. For example, a decrease in the number of unique Bluetooth devices encountered or a decline in WiFi scans could indicate reduced social interaction or maybe a person is spending time somewhere alone [BNW+18]. Similarly, the persistence of interactions with specific devices might reveal the strength and regularity of social ties.

- **Bluetooth or WiFi Scans:** The count of scans represents the total number of times Bluetooth devices or WiFi access points are detected during a specific time segment. A higher number of scans may indicate more frequent detection of nearby devices that may show being present in a socially active environment. However, it is important to interpret these findings cautiously, as decreased scans could also be due to other factors which could be unrelated to social interaction, such as device settings or surroundings conditions.
- **Unique Device Count:** For both Bluetooth and WiFi, the number of unique devices or access points sensed in a given time segment is recorded which is find through MAC addresses. This metric could tell the diversity of potential interactions and suggests the individual's movement across different social settings.
- **Mean and Standard Deviation of Scans:** By calculating the mean and

standard deviation of scans for sensed devices can help in understanding the regularity and consistency of interactions.

3.4.8 Sleep Features

Sleep patterns are important factors to determine overall health and well-being. Disturbances and rregularities in sleep patterns are often associated with various psychological conditions, including loneliness [ML16, GWR+20]. Sleep features are computed for a single epoch, which is per day. Below are some key features related to sleep:

- **Sleep timing and duration:** The times when individuals first wake up and their last bedtimes are tracked to find their sleep schedules. Similarly, the total time spent while sleeping is recorded as sleep duration.
- **Sleep efficiency:** This is measured by the ratio of the actual time spent asleep compared to the total time spent in bed. A higher ratio means better sleep quality, as it means a greater proportion of time in bed is spent sleeping.
- **Time in bed:** This includes both sleeping and awake time while in bed.

3.5 Data Cleaning

Data cleaning is an important step for performing a meaningful analysis. This involved applying different operations to the data to improve its quality and reliability before using this data for machine learning model training.

- Participants Selection: For our analysis, we included only those students who had completed the post-study loneliness survey. It was to ensure that our analysis was based on a comprehensive and reliable measure of each participant's perceived loneliness at the end of the study period. This approach resulted in a final sample of 41 students out of 48 for the StudentLife dataset and 205 out of 218 for the GLOBEM dataset. Our rationale was based on the idea that post study measures more accurately reflect the student's actual feelings of loneliness during the semester. We considered that these feelings might be influenced by academic pressures, such as exams or project deadlines, as well as the overall social relationships of students throughout the semester. This makes post-study loneliness measures a more valid representation of the student's mental state throughout the academic term. While this approach provides an end of study assessment, we acknowledge that it may not capture fluctuations in loneliness levels throughout the study period.

- Outlier Removal: It is important to identify and remove outliers from the numerical features in the data since they can influence on different analyzes and model training. We used the Interquartile Range (IQR) method for this due to its suitability for sensor data, which often contains non-normally distributed values and sometimes extreme readings. We performed outlier removal separately per student data since sensor data can vary significantly between individuals. We chose the commonly used 1.5 * IQR threshold as it has a balance between removing extreme values that could skew analysis and retaining potentially meaningful variations in the data. The IQR method works as follows:
  - Calculate Q1 (25th percentile) and Q3 (75th percentile) for each feature
  - Compute IQR = Q3 - Q1
  - Define outliers as any data points falling outside the range [Q1 - 1.5 * IQR, Q3 + 1.5 * IQR]
- Feature Filtering: We applied a feature filtering step to reduce noise in data. First, we excluded features with zero variance. These are the features which provide no discriminative power for our models. As features with constant values across all samples do not contribute to the model's learning process and can potentially introduce unnecessary complexity. Then we removed features with more than 30% missing data as this is a common threshold in passive sensing based studies. This threshold was chosen as a balance between retaining potentially valuable information and ensuring the reliability of our features. High levels of missing data can lead to biased or unreliable models, and missing imputation techniques become less reliable with excessive missing values. We also assessed multicollinearity among the features using Variance Inflation Factors (VIFs) and subsequently removed any features with a VIF exceeding 5 as this threshold indicates a high level of multicollinearity that could compromise model stability and reliability. In cases where multiple features were highly correlated, we retained the feature with the highest predictive power based on preliminary model testing. We applied these filtering criteria to the training set only and then used the

resulting feature set for both training and test sets.

- Missing Value Imputation: After filtering out features with excessive missing data, we addressed the remaining missing values in our dataset. We used the median for numerical features to impute missing values and the mode for the categorical features. We applied this in train sets and used the median and mode derived from train set to impute into validation and test sets. While other approaches could also be considered like ffill and bfill but sensor data often has variability and irregular patterns, which makes these methods less suitable due to their potential to introduce bias from preceding or succeeding values. That's why using the median and mode was a more generalized approach since it maintains dataset integrity and stability for models and also reduces the risk of overfitting to specific temporal patterns.

3.6 Models Training and Evaluation

Throughout this thesis, we have performed different statistical analyses and trained different machine learning models for loneliness detection in different contexts and research questions. Here, we will present the generic approach of the trained models, while specific details will be presented in each chapter. We used UCLA response thresholds as we discussed above to divide loneliness into 2 categories for binary classification problem (lonely/not lonely). We used these class categories to label our data. We chose binary classification due to the relatively small size of the datasets and the limited 10-week period of data collection. Since loneliness levels change very slowly, categorizing it into multiple levels or treating it as a continuous variable would have been challenging with the limited data available. We mainly use logistic regression, k-nearest neighbours (KNN), support vector machine, random forest, and XGBoost algorithms for this classification.

We chose these ML algorithms because they are widely used in supervised classification, easy to train, and interpretable [GCRN+18]. Moreover, these ML algorithms have also been used in previous studies that we discussed in Chapter 2. We used the mutual information (MI) method with cross-validation to select optimal features for our models. Mutual information measures how much a feature tells us about the target variable. Higher MI scores indicate more informative features. It works by first ranking features in descending order based on their MI scores and then calculating the cumulative sum of these scores. Features are then selected until their cumulative MI reaches a 95% threshold of the total MI.

This approach helps to ensure that the selected features collectively explain a substantial portion of the variance in loneliness while avoiding redundancy and potential overfitting.

To handle biases in the model due to class imbalance (lonely/not lonely), we used the synthetic minority oversampling method (SMOTE) [CBHK02]. We implemented SMOTE for each fold of cross-validation on the training data. This method creates synthetic examples of the minority class without discarding existing data or introducing excessive duplication. It leads to a more balanced training set for our classification models. This approach helps the model to get trained equally for both classes and reduce the bias toward the majority class to improve the overall performance of the classification models [EA19]. For hyperparameter optimization, we used grid search to select the best parameters. It works by defining a range of hyperparameter values for each model. These ranges were decided based on pilot runs. Then, this systematically tests model performance for each combination of hyperparameters. Finally, it selects the combination that provided the best cross-validated performance. We used both nested cross-validation for hyperparameter tuning and k-fold cross-validation for performance assessment, depending on the specific requirements of each chapter. Detailed explanations of the cross-validation procedures used in each chapter are provided in their respective methodology sections. The performance metrics used for evaluation included accuracy, precision, sensitivity, and F1 Score.

1. Accuracy: Provides an overall evaluation of the model's ability to correctly classify students as 'lonely' or 'not lonely'. It is calculated as the proportion of correct predictions (true positives and true negatives) out of the total number of cases examined.

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \tag{3.1}$$

   where TP is True Positives, TN is True Negatives, FP is False Positives, and FN is False Negatives.

2. Precision: This metric focuses on the model's accuracy in identifying individuals who are truly lonely. It is the proportion of correctly identified lonely individuals (true positives) among all those classified as 'lonely'. High precision is important in loneliness detection to minimize false alarms.

$$\text{Precision} = \frac{TP}{TP + FP} \tag{3.2}$$

3. Sensitivity (Recall): This metric measures the model's ability to identify all truly lonely individuals within the dataset. It is the proportion of correctly identified lonely individuals (true positives) out of all actual lonely individuals. High sensitivity is crucial in ensuring that no lonely individuals are missed, as overlooking them could result in a lack of timely support and intervention.

   Sensitivity has greater relevance than specificity in the context of loneliness detection as the primary goal is to identify individuals experiencing loneliness correctly. Missing a lonely individual (false negative) can have significant implications which makes it crucial that the model captures as many true positives as possible. Therefore, we focused on reporting sensitivity along with precision and accuracy to provide a balanced view of the model's performance in identifying lonely individuals. Specificity, on the other hand, measures the model's ability to correctly identify those who are not lonely (true negatives). While it is important to minimize false positives, specificity is less critical in this context, where the priority is on avoiding missed cases of loneliness.

$$\text{Sensitivity} = \frac{TP}{TP + FN} \tag{3.3}$$

4. F1 Score: This provides a balanced evaluation of the model by considering both precision and sensitivity as it provides the harmonic mean of both metrics. It shows the model's ability to accurately detect loneliness while minimizing both false positives and false negatives.

$$\text{F1 Score} = 2 \times \frac{\text{Precision} \times \text{Sensitivity}}{\text{Precision} + \text{Sensitivity}} \tag{3.4}$$

Chapter 4

Personalizing Loneliness Detection Through Behavioural Grouping of Passive Sensing Data From College Students

4.1 Overview

As we discussed in Chapter 1, loneliness is a serious global issue specifically among younger population. While Chapter 2 highlighted the potential of passive sensing for loneliness detection in the existing literature but traditional one-size-fits-all approaches may overlook the diverse ways loneliness manifests in student populations. Current passive sensing-based methods for detecting loneliness use generalized models trained with all available data at once. In this method, the model learns broader patterns that are common among observations, which are then used for loneliness detection. Loneliness is a subjective experience, and its behavioural manifestations can vary among individuals based on their personality type and social contexts. Generalized models may overlook these subtle variations by focusing on broader patterns that are common across the entire population. However, by identifying subgroups of individuals with similar behavioural patterns, we can develop more targeted and personalized models that better reflect the unique ways loneliness is expressed in different groups.

We hypothesize that by identifying and analyzing subgroups of college students with similar behavioural patterns, we can capture more nuanced and specific behaviours related to loneliness that generalized models trained on the entire student population may overlook. Moreover, it is also important to analyze group dynamics over time to monitor changes and track the evolution of groups over time. This helps in identifying behavioural patterns and trends that are not apparent through traditional methods.

This chapter presents an empirical evaluation of the effectiveness of group-based loneliness classification models in college students using passive sensing. We used incremental clustering to identify groups of students with similar behaviours and to monitor changes in these groups over time. Subsequently, we developed tailored loneliness classification models for each group to compare group-based models with generalized models. In addition, this chapter also provides group-specific behavioural patterns within students. While clustering approaches have been utilized in various fields of studies [XNG14, HMY22], their application to the understanding and classification of loneliness, particularly in the context of college students with passive sensing data is limited. All existing research on loneliness detection using passive sensing is based on generalized machine learning models for all available data that does not consider inter-subject differences or behavioural variations across groups. Moreover, existing studies have not analyzed the behavioural shifts of students in relation to loneliness during a study period. Our approach could lead to the development of more targeted and personalized interventions that consider the unique behaviours and experiences of different student subgroups. The key research questions we have explored in this chapter are:

1. Can clustering methods identify distinct behavioural groups among college students using passive sensing data, and how do these groups evolve over time?

2. How effective are group-based loneliness classification models compared to generalized models among college students using passive sensing data?

4.2 Methodology

4.2.1 Finding Behavioural Groups from Student Data

This study used the StudentLife dataset. We have mainly used an incremental clustering approach to capture the dynamic nature of student behaviours. Unlike traditional clustering methods, incremental clustering can process data as it arrives which allows for the identification of evolving behavioural patterns and timely updates to the clustering model. Since students' behaviours can vary over time due to academic pressures, social interactions, and personal circumstances, incremental clustering makes sense and can reflect real-time changes in student behaviour. To establish a baseline for comparison and to evaluate the clustering results' stability, we used multiple clustering algorithms to gain a broader understanding of the grouping in the dataset. We used weekly aggregated data as input for clustering methods.

K-means is a classic partitioning algorithm that identifies clusters with well-defined centroids. [SY20]. It is commonly used in behavioural analysis due to its simplicity and scalability. OPTICS is a density-based method used in handling datasets with varying densities [ABKS99]. PAM is similar to K-means that finds representative data points (medoids) as cluster centers, which helps in adding potential robustness against outliers often present in behavioural datasets [VdLPB03]. IDBSCAN is also a density-based algorithm with an incremental version of the usual DBSCAN, and it works in identifying clusters of varying densities incrementally and handling noise within the data [EKS+98]. This selection of both partitioning (K-means and PAM) and density-based (OPTICS and IDBSCAN) algorithms was to capture the diverse nature of student behavioral patterns. K-means and PAM are effective when clusters are well-separated and spherical which may align with potential distinct groups of students based on factors like academic performance or social engagement. However, recognizing that student behaviours may not always fit into such categories, OPTICS and IDBSCAN can provide the flexibility to identify clusters with varying densities and shapes which might be overlooked by traditional methods.

Before applying the clustering techniques, we first normalized the weekly data using min-max scaling to ensure all features were on the same scale. Then, we applied Principal Component Analysis (PCA) to reduce the dimensionality of the data [JC16]. PCA combines the original variables (behavioural features) into a smaller set of new variables called principal components. These principal components are like condensed summaries of the original data that capture the most significant variations in the behavioural patterns while using fewer dimensions. We determined the optimal number of principal components through the explained variance threshold, scree plot analysis, and empirical experimentation to evaluate how the number of components affects clustering quality results.

4.2.1.1 Determining the Optimal Number of Clusters

For K-means and PAM algorithms, we must define the number of potential clusters beforehand. We used the elbow method to make this selection [JN13]. This method works by iteratively fitting the clustering model with different numbers of clusters (K values). For each K, a measure of clustering quality (within-cluster sum of squared errors) is calculated. By plotting these quality scores against the corresponding K values, the elbow point on the graph shows the K value beyond which adding more clusters does not significantly impact. The goal is to achieve a compact and well-separated clustering solution without unnecessarily complicating the model. We initially determined the optimal number of clusters using Week 1 data and then re-evaluated this on Week 5 based on cluster quality results to ensure the clusters remain relevant as more data accumulates.

4.2.1.2 Evaluating Cluster Quality

We used two evaluation metrics to assess the quality of clusters: the Silhouette score [RS+22] and the Xie-Beni Index [XB91]. The Silhouette score evaluates the consistency of clusters. It measures how well each data point fits within its assigned cluster compared to its similarity to points in neighbouring clusters. It provides a value between -1 and 1, where a high value shows that the object is well matched to its own cluster and poorly matched to neighbouring clusters. This score validates the consistency within clusters and the differentiation between them which tells whether the data points have been grouped into meaningful clusters or not. The Silhouette score is particularly useful for confirming the appropriateness of the number of clusters chosen and it ensures that each cluster represents a distinct group with similarities.

The Xie-Beni index provides a combined assessment of cluster compactness and separation. Compactness means how closely data points are grouped within a cluster, while separation measures the distance between distinct clusters. A lower Xie-Beni Index means a better clustering structure with tight and well-separated clusters. This index is useful for confirming that the clusters are distinct and do not overlap significantly. We calculated these scores weekly to continuously assess clusters quality.

4.2.1.3 IDBSCAN Algorithm

After evaluating the clustering results produced by different algorithms in Figure 4.1a and Figure 4.1b, we selected IDBSCAN for further analysis. It is an incremental variation of the density-based DBSCAN algorithm that identifies clusters based on the density of data points in a region. This makes it suitable for discovering clusters of arbitrary shapes and handling outliers, unlike traditional distance-based methods like k-means [XT15]. It is designed to adapt to evolving data distributions. This aspect of the algorithm helps to update and refine cluster assignments as new data becomes available without retraining the whole model with all the data available. This incremental approach aligns with the evolving nature of student behaviours captured through passive sensing data which provides a more accurate and up-to-date reflection of students' loneliness-related behavioural patterns. Upon the arrival of new data, the algorithm calculates the mean distance between each new item and the core objects of existing clusters. If a new data item is within the specified radius (eps) of any cluster core and the number of points in the cluster satisfies the minimum required threshold (minPts), the item is added directly to the nearest cluster. Otherwise, if the item does not meet these criteria, it is initially marked as an outlier. Outliers that accumulate to meet the minPts threshold can subsequently form new clusters.

To find the optimal values for the eps (distance threshold) and minPts (minimum points) parameters, we used a two step approach. First, we conducted pilot runs with different combinations of eps and minPts values to assess cluster quality using the Silhouette score and Xie-Beni index. These initial experiments helped us to identify suitable ranges for both parameters. Then we used GridSearchCV with cross-validation to find the best values through the ranges identified. This involved systematically testing various combinations of eps and minPts values found through pilot runs and evaluating their performance for cluster quality. The combination that resulted in the highest average silhouette score across the cross-validation folds was then selected as the optimal parameter.

4.2.2 Loneliness Detection for each Group

After identifying behavioural subgroups through weekly incremental clustering, we implemented a personalized approach to loneliness detection. This personalized approach acknowledges that the relationship between behavioural patterns and loneliness may vary across different student subgroups. We trained four binary classifiers for each subgroup; logistic regression, random forest, support vector machines (SVMs), and XGBoost. To evaluate and select the best model for each group, we used 10-fold stratified nested cross-validation with a 3-fold inner loop for hyperparameters tuning. It is important to note that leave-one-subject-out cross-validation (LOSO) was not applied in this chapter. LOSO excludes one participant at a time for model testing which provides a more robust evaluation of model generalizability. Due to the small size of the dataset and the group-based approach used in this chapter, applying LOSO would have been challenging, as removing one participant from small groups could lead to overly sparse training sets and impact model stability. That is why, we relied on standard cross-validation for evaluating model performance.

We handled missing values, feature scaling, feature selection, applied SMOTE to address class imbalance in the inner fold of the cross-validation. We applied the same scaling parameters, imputation values and selected features derived from the training set to the corresponding validation and test sets. We selected the best model in the inner loop using F1-macro. The final performance metrics used for evaluation included accuracy, precision, sensitivity, and F1 Score.

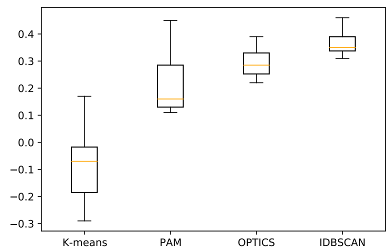
4.2.3 Loneliness Detection using all Students Data

To evaluate the effectiveness of our personalized, group-based approach, we compared its performance to generic models trained on the combined dataset without any grouping. For the generic models, we used the same classification algorithms as in the group-based approach: logistic regression, random forest, support vector machines (SVMs), and XGBoost. We again used the same nested cross-validation and other steps with 10 outer folds and 3 inner folds for hyperparameter tuning and model selection. The final performance metrics used for evaluating both the generic and group-based models included accuracy, precision, sensitivity, and F1 Score.

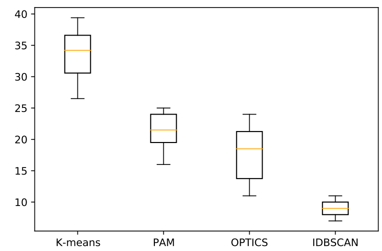
4.3 Results

4.3.1 Evaluating Cluster Quality

To assess the effectiveness of each clustering algorithm, we evaluated cluster quality using the Silhouette score and the Xie-Beni Index. Figure 4.1a shows the Silhouette scores for each of the algorithms used in a boxplot form. Based on the results, IDBSCAN has the highest Silhouette score which shows that data is properly combined in groups for this algorithm and that the inter-cluster and intra-cluster distances are excellent for those groups. Similarly, Figure 4.1b shows the Xie-Beni Index scores for different algorithms in which IDBSCAN has the lowest score, which means clustering results are better for IDBSCAN as compared to other methods in terms of compactness and separation of the clusters. The idea behind the Xie-Beni Index is to minimize the distance between points within a cluster while maximizing the distance between different clusters. OPTICS and PAM also showed reasonable results which indicates a some degree of cluster structure. IDBSCAN's and OPTICS ability to find clusters of arbitrary shapes and their adaptability to varying densities within the data likely contributed to their superior performance.



(a) Boxplot comparing Silhouette scores



(b) Boxplot comparing the Xie-Beni Index

Figure 4.1: Clustering Algorithms Comparison

4.3.2 Temporal Evolution of Student Groups and Behavioural Profiles

To examine the dynamic nature of student behaviours over the 10-week period, we applied the clustering in an incremental way. Data from each subsequent week was integrated into the existing model. The aim was to track behavioural changes and the evolution of groups throughout the study period of 10 weeks. To monitor changes in students' group memberships over time, each student was assigned to the cluster where the majority of their weekly data points belonged for that specific week. This means that students are not fixed in one group for the entire study but can move between groups over time that reflects shifts in their behavioural patterns. By the end of Week 10, it is possible for a student's data points to belong to multiple groups across different weeks. The purpose of these weekly group assignments is to provide a more dynamic understanding of how behavioural patterns evolve over time and to show the changing nature of group membership over the 10-week period.

Figure 4.2, shows the evolution of student groups over the 10-week period. The x-axis represents each week while the y-axis shows the number of groups and number of students per group. It shows the stability of groups and transitions of

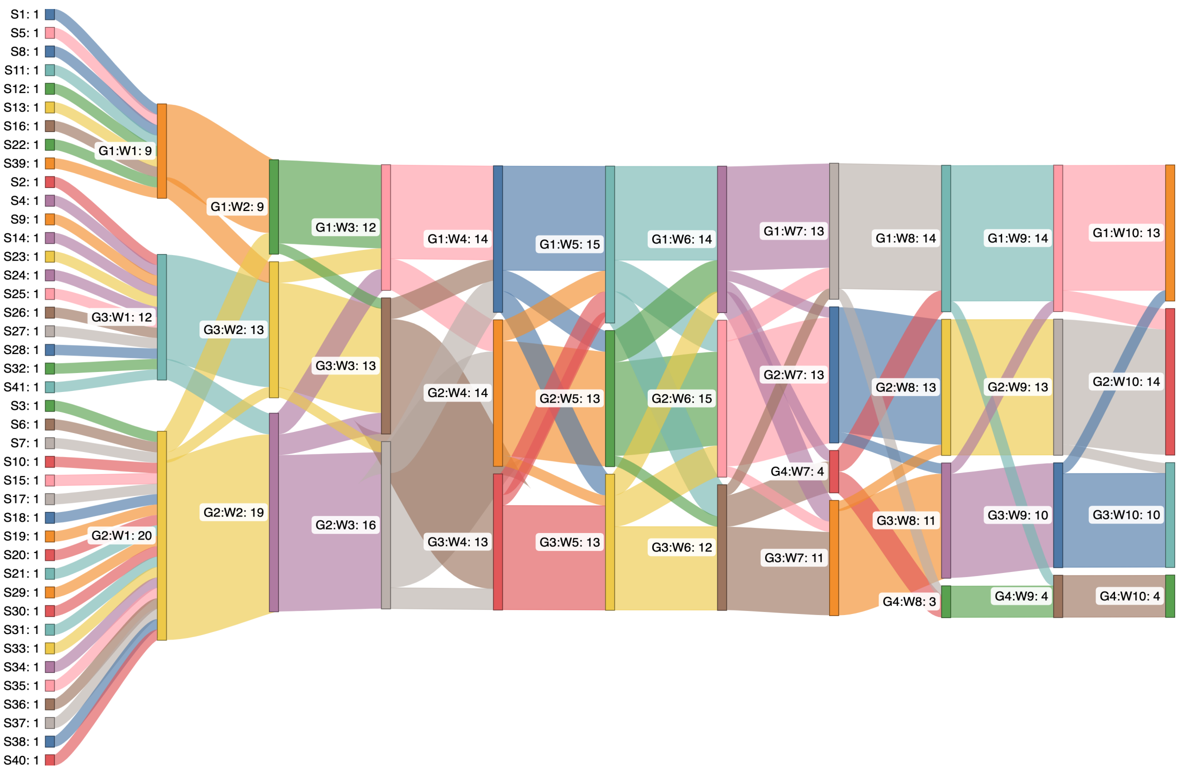
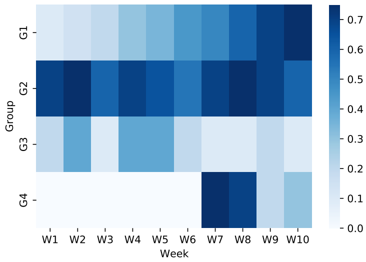


Figure 4.2: Visualization of Students Temporal Behavioural Group Dynamics over a 10-Week Period (G1:W1:9 denotes Group 1 in Week 1 with 9 students)

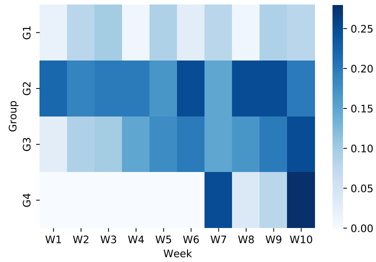
students within the group dynamics. Initial analysis from weeks 1 to 6 show a relatively stable grouping with three distinct behavioural groups. These groups are labelled G1, G2, and G3 in the figure, with the number of students in each group indicated beside the group label (e.g., G1:W1:9 denotes Group 1 in Week 1 with 9 students). A change occurred from Week 7 onwards, with the model identifying four distinct groups (labelled G1, G2, G3, and G4). This shows a potential change in behavioural patterns across the student population around this time. It is important to note that the group sizes also show fluctuations throughout the ten weeks which highlights the dynamic nature of the group members.

4.3.2.1 Group Behavioural Profiles Analysis

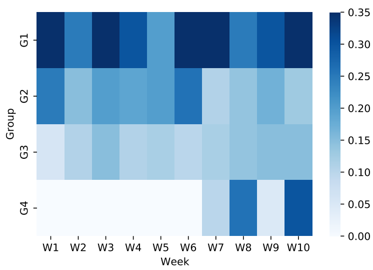
To get a deeper understanding of the distinct behavioural patterns represented by each group, we analyzed average group features on a weekly basis. This aggregation gives a clear and interpretable view of the dominant behavioural tendencies within groups over time. We aimed to identify the main behavioural patterns that differentiate each group by focusing on key behavioural metrics (activity duration, conversation duration, phone usage duration, and location transitions). Figures 4.3a, 4.3b, 4.3c and 4.3d presents a visual representation of how key behavioural features vary across the student groups over the 10-week study period.



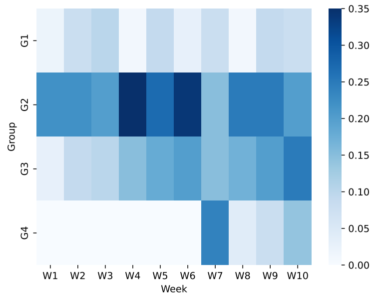
(a) Activity duration



(b) Conversation duration



(c) Phone usage duration



(d) Location Transitions

Figure 4.3: Average Weekly Behavioural Patterns (normalised) for 4 groups over a 10-week period

Activity Patterns

Figure 4.3a presents a heatmap that shows the activity duration patterns for each group throughout the study period of 10 weeks. This provides a comparison of the duration of physical activity. Groups 1 and 3 had overall shorter duration of activity as compared to Group 2. An observable pattern is the gradual increase in activity duration for Group 1 which shows a change in their activity level over a period of time. Group 4 had higher levels for weeks 7 and 8, while they decreased for the last 2 weeks.

Conversation Patterns

Figure 4.3b shows the weekly trends of conversation duration across different groups. This highlights the subtle variations in social interactions. Group 1 continuously shows shorter conversation durations, whereas Group 2's reaches its highest point around week 6 and again in the final weeks. The conversation durations of Group 3 gradually increased. Group 4 has a less consistent pattern, which shows fluctuations in their social interactions.

Phone Usage Patterns

The analysis of phone usage duration patterns in Figure 4.3c highlights Group 1 as the most engaged phone usage throughout the study, while Group 3 consistently shows lower engagement levels. Group 2's phone usage fluctuates between moderate and low, while Group 4's shows a low trend.

Location Transition Patterns

Figure 4.3d shows the weekly patterns for average location transitions for 4 groups. The number of location transitions represents the number of movements between different places. Group 1 has the lowest location changes as compared to other groups. Group 2 shows a moderate to high trend, while Groups 3 and 4 have a low trend.

4.3.3 Loneliness Detection Using all Students Data

This section presents the performance of four binary classification models, specifically Logistic Regression, Random Forest, Support Vector Machine (SVM), and XGBoost, which are trained on all the data at once. This performance of the generalized models for the 'lonely' class is summarized in Table 4.1.

Table 4.1: Performance of centralised machine learning models classifying loneliness using a subset from the StudentLife dataset.

Model	Accuracy,%	Precision,%	Sensitivity,%	F1 Score,%
XGBoost	81.11	91.48	86.07	88.50
RF	82.32	86.07	82.36	84.14
SVM	76.50	75.83	82.07	79.16
LR	72.04	67.16	78.47	72.35

We can see that ensemble-based classifiers (XGBoost and Random Forest) performed overall well as compared to SVM and LR. Among these generic models, XGBoost achieved the highest overall performance (except accuracy) with an accuracy of 81.11%, a precision of 91.48%, a sensitivity of 86.07%, and an F1 score of 88.50%. This shows that XGBoost was particularly effective at correctly predicting loneliness when it identified it (high precision) while also successfully detecting a large proportion of all actual loneliness cases in the dataset (high sensitivity). The Random Forest models followed closely in performance, with the highest accuracy of 82.32%. However, precision, sensitivity, and F1 scores were a bit lower, at 86.07%, 82.36%, and 84.14% respectively. SVMs and logistic regression performed less effectively, with accuracies of 76.50% and 72.04%, respectively.

4.3.4 Group Based Loneliness Detection

To directly assess the impact of our personalized approach, we trained loneliness classification models specifically for each behavioural group identified through our clustering process. Over the 10-week period, we applied incremental clustering weekly to assign each student's data points to one of the evolving behavioural groups. Instead of assigning a single group to each student for the entire period, each data point was independently assigned to a group based on that week's clustering results. For each of the four behavioural groups identified, a distinct binary classification model was trained. Each model utilized all data points assigned to its respective group across the 10 weeks. Since a student's data points could belong to different groups in different weeks based on clustering results, their data contributed to multiple group-specific models at final week 10. This approach helps in accounting for behavioural fluidity to ensure that the models reflect the diversity of experiences across the student population.

Figure 4.4 shows sample training and test records for two behavioural groups identified through weekly clustering. For example, we can see that Student A, who has a consistent positive (lonely) label, appears in G1 in Week 3 alongside Student C, who also has a positive label, but switches to G2 in Week 4, where they are clustered with Student D, who has a negative (not lonely) label. For training and testing each group-specific model, we applied 10-fold cross-validation within each group.

There are a few limitations associated with this approach of using a single post-study loneliness label for all weekly records. Since the dataset only provides one loneliness score per student collected at the end of the study, we had to apply this final score as a label to all records for that student each week. This limitation restricts the model's ability to capture short-term changes in loneliness that might occur week-to-week. Ideally, if we had weekly loneliness responses, they would provide a more granular and accurate view of the fluctuations in loneliness over time and we could label each week's records based on weekly self-reports. Moreover, using weekly data with a final, static loneliness label may have introduced noise, especially in weeks where a student's behaviour differs from the

Sample Records For Group 1:

Student	Week	Features	Label
A	3	…	1
C	3	…	1
…	…	…	…

Sample Records For Group 2:

Student	Week	Features	Label
A	4	…	1
D	4	…	0
…	…	…	…

Figure 4.4: Sample Data Records For Two Different Groups

overall loneliness state indicated by the end-of-study score. This discrepancy between weekly behaviour and final labels can potentially confuse the model leading to lower predictive performance. To address these limitations in future work, we have collected weekly loneliness survey responses in our own collected new dataset. By obtaining weekly labels, we aim to better capture the nuances and temporal dynamics of loneliness, which will help create more accurate and contextually relevant models that reflect the weekly experiences of participants.

Table 4.2 presents the performance metrics of these group-based models. For the majority of the groups (Groups 1, 2, and 3), the detection models show significantly improved performance compared to the generic models trained on the entire dataset. Across most groups, XGBoost emerges as the top-performing algorithm. In G1, XGBoost achieved an accuracy of 88.5%, precision of 94%, sensitivity of 91.5%, and an F1 score of 92.5%. Its performance slightly improved in G2, with an accuracy of 89.5%, precision of 95.5%, sensitivity of 90.5%, and an F1 score of 93.7%. For G3, XGBoost maintained strong performance with an accuracy of 86.5%, precision of 94.3%, sensitivity of 88%, and an F1 score of 92.3%. Random Forest ranked second in performance across G1, G2, and G3, with F1 scores of 89.4%, 89.3%, and 87% respectively. SVM showed moderate performance, with F1 scores ranging from 83.4% to 86% across these groups. Logistic regression performed poorly as compared to other algorithms. All models showed a significant drop in performance for G4. XGBoost, while still outperforming other models, reduced its accuracy to 61.5% and its F1 score to 48.5%. The other models also showed declines in performance with F1 scores ranging from 42.5% to 45% for G4.

Table 4.2: Performance of behavioural group based machine learning models classifying loneliness using a subset from the StudentLife dataset.

Algo	G1				G2				G3				G4			
	Acc.	Prec.	Sens.	F1.	Acc.	Prec.	Sens.	F1.	Acc.	Prec.	Sens.	F1.	Acc.	Prec.	Sens.	F1.
XGBoost	88.5	94	91.5	92.5	89.5	95.5	90.5	93.7	86.5	94.3	88.0	92.3	61.5	55.5	44.2	48.5
RF	87.1	92.0	87.4	89.4	89.0	93.4	85.5	89.3	85.5	92.4	82.1	87.0	55.4	48.5	41.3	45.0
SVM	81.4	84.5	88.0	86.0	78.5	82.5	86.3	84.4	82.3	80.5	85.0	83.4	40.4	46.5	42.3	44.5
LR	74.5	70.3	83.7	76.4	76.3	72.6	79.4	76.0	77.0	73.4	80.0	76.7	41.5	40.0	44.2	42.5

Algo:Binary Classification Algorithm; G1: Group 1; G2:Group 2; G3:Group 3
G4:Group 4; Acc:Accuracy(%); Prec:Precision(%); Sens:Sensitivity(%); F1:F1 Score(%)
RF:Random Forest; SVM:Support Vector Machine; LR:Logistic Regression

4.3.5 Aggregated Performance of Personalized Models

To assess the overall effectiveness of group-based approach for loneliness detection compared to the generic models, we calculated the aggregated performance metrics for the personalized models across all groups. We used a weighted average based on the number of data points in each group to ensure that the aggregated metrics accurately reflect the contributions of each group. Table 4.3 presents the performance metrics for aggregated personalised models.

Table 4.3: Aggregated Performance Metrics of Personalized Models Across All Groups

Model	**Accuracy,%**	**Precision,%**	**Sensitivity,%**	**F1 Score,%**
XGBoost	85.72	90.83	88.69	89.78
RF	84.33	88.33	82.96	85.58
SVM	76.62	79.13	82.24	80.77
LR	72.49	68.90	77.46	73.02

The personalized XGBoost model outperformed the generic model in terms of accuracy (85.72% vs. 81.11%) and F1 Score (88.57% vs. 89.78%). Similarly, the personalized Random Forest model showed higher accuracy (84.33% vs. 82.32%) and F1 Score (85.58% vs. 84.14%), while sensitivity was almost similar. The performance for SVM and logistic regression was almost same.

4.4 Discussion

4.4.1 Key Findings

This chapter investigated the potential of using unsupervised learning to create a personalized approach for student loneliness detection. The clustering approach effectively identified various behavioural groups among the student population. This shows that loneliness is not a uniform experience and that it is subjective in nature. Students show diverse behavioural patterns that may indicate different underlying causes or expressions of loneliness. This finding aligns with research highlighting the subjective nature of loneliness and its varying manifestations across individuals [QVH+15, CP08]. Furthermore, the personalized detection models achieved significantly improved performance compared to generic models trained on the entire dataset. This finding highlights the limitations of the one-size-fits-all approach. Generic models may struggle to capture the nuances of how loneliness manifests within specific behavioural subgroups. By tailoring detection models to the unique patterns of each group, we were able to achieve more accurate identification of students experiencing loneliness. Our dynamic temporal analysis shows the potential for tracking changes in student behaviour and group participants over time. This finding suggests that loneliness may not be a static state and students' behavioural patterns may shift throughout the semester. These shifts could be linked to external factors like academic pressure, social connections, or changes in mood [CH03]. Identifying such transitions provides significant opportunities for early loneliness detection and timely intervention. If a student's behavioural patterns indicate a move towards a group associated with higher loneliness risk, proactive support could be provided before loneliness significantly impacts their well-being.

4.4.2 Observations and Implications

A close examination of the behavioural profiles across the identified groups shows important differences in how students' behaviours varied across groups. Groups 1 and 3 show relatively low activity levels and low location changes. This pattern aligns with existing literature on loneliness, which suggests that reduced physical engagement with the outside world and potential reliance on technology to fulfill social needs [CH09, PSS+17]. Reduced physical activity could indicate in-person social withdrawal which is reported as a common symptom of loneliness in literature [CH09]. Furthermore, the high phone usage in G1 might indicate a substitution of digital interactions for in-person social engagement which is a phenomenon observed in previous studies linking loneliness to increased technology use [PSS+17]. However, it is important to acknowledge that the relationship between technology use and loneliness is complex, with technology potentially serving as both a tool for social connectivity and a contributor to isolation depending on the quality and nature of the interactions [NNC18].

In contrast, Group 2 shows a significantly different profile, with high activity levels, frequent conversations, and moderate phone usage. This pattern shows a moderate degree of physical and social engagement, likely reflecting an active social life with diverse interactions. Students within this group may be less vulnerable to experiencing loneliness, as individuals with diverse and meaningful social connections tend to have better emotional regulation and coping mechanisms, reducing their susceptibility to loneliness [MCHC11]. However, it is important to acknowledge that these are speculative interpretations, and further research is needed to establish a definitive causal relationship between these behavioural patterns and loneliness. It is possible that other factors not captured in the dataset could influence both behaviour and loneliness experiences.

The emergence of a new group during Week 7 is an interesting observation that shows the importance of dynamic analysis. Existing literature emphasizes the fluctuating nature of loneliness and its susceptibility to various influences over time [VLS+13]. Several factors are likely to have had an influence on this finding. Expanding the sample by including more weeks of data might have uncovered a wider and more nuanced range of behavioural patterns. The increased complexity probably caused the model to identify a distinct group that was not as evident in the first set of data. Additionally, it is possible that underlying shifts in student behaviours occurred around Week 7. These shifts could be due to many reasons like evolving social dynamics, changes in academic workload, or other external events. This aligns with research highlighting the impact of life events and social dynamics impact on evolving loneliness experiences [CH03]. These shifts could have created newly identifiable differences between student subgroups. Furthermore, the model itself have evolved over time, becoming more sensitive to subtle behavioural variations and better able to distinguish distinct groups within the dataset.

The ML model results have shown the superior performance of ensemble methods (XGBoost and Random Forest) over traditional methods like SVM and Logistic Regression across all behavioural groups. This performance difference is likely due to the inherent characteristics of ensemble methods that make them well-suited for complex, non-linear relationships which are often present in behavioural data. While our personalized detection models performed well for most groups, the results for Group 4 raise important concerns. This group shows the lowest classification performance for loneliness detection which highlights challenges of effectively classifying students within this subgroup. The relatively small data size within Group 4 and due to its late emergence (week 8) may have made it challenging for the model to learn the subtle patterns associated with loneliness for this specific group. This performance drop in Group 4 affected the aggregated metrics as well and highlighted limitations in our approach for smaller groups. This indicates that while personalization enhances model performance for most groups, it may not be as effective for groups with limited data. However, this needs further investigation to find the underlying reasons for poor performance. This shows the complexity of loneliness detection and the importance of considering group heterogeneity when developing and evaluating predictive models. Simply relying on generic models trained on the entire dataset may fail to capture these nuances and could lead to less accurate predictions and missed opportunities for timely intervention.

However, one of the important factors to consider is that the absence of leave-one-subject-out cross-validation (LOSO) may have influenced these results. Without LOSO, the models may be more prone to overfitting to the specific individuals present in the training set. This could explain why the models perform better within smaller groups, where behavioural patterns are more homogeneous and fewer students are present. The use of LOSO could help assess the generalizability of these models more effectively by ensuring that each student is tested in isolation from the training set.

4.5 Conclusions and Limitations

This chapter investigated the potential of using unsupervised learning to create a personalized approach for student loneliness detection and the findings directly address the key research questions:

- Our results show that clustering methods can identify distinct behavioural groups among college students using passive sensing data. Moreover, the temporal analysis showed how these groups evolve over time. This highlights the fluid nature of student behaviour and the importance of longitudinal tracking.
- The superior performance of personalized, group-based loneliness classification models compared to generic models trained on the entire dataset shows the value of our approach. This finding empirically validates the hypothesis that accounting for inter-subject variability through behavioural grouping can lead to more accurate and effective loneliness detection. By identifying subgroups with similar behavioural patterns, we can develop more targeted and personalized detection models that better reflect the unique ways lone-

liness is expressed in different groups.

Even though these findings show improved performance for group-based models for loneliness detection, there are some limitations that need to be addressed. These limitations may lead to new research questions for exploration as mentioned in future work (see Chapter 8 for details).

- The study is based on a specific cohort of 41 college students which limits the generalizability of the findings to other populations or age groups. College students often share similar daily routines and other psychological pressures, which reduces the generalizability of our results.
- Another important limitation of our study is in the reliance on the single StudentLife dataset. While this dataset provides a good amount of behavioural data, it does not capture the full spectrum of factors that can influence an individual's experience of loneliness. Other important aspects, such as demographics data, family history, personal relationships, major life events, and mental health history are not included in this data. These factors play a significant role in shaping how loneliness manifests and could interact with behavioural patterns [BG88].
- While our group-based models show improved performance, future work should aim to include leave-one-subject-out cross-validation (LOSO). Applying LOSO will provide a more rigorous assessment of the model's ability to generalize to new individuals by isolating participants from the training set during evaluation.

Chapter 5

Dynamic and Personalized Loneliness Detection Using Passive Sensing

5.1 Overview

Chapter 4 presented the potential of personalized, dynamic behavioural analysis for improved loneliness detection within student populations. We have effectively identified distinct behavioural subgroups showing that loneliness can manifest in diverse ways among students. Furthermore, tailored models outperformed generic models trained on the entire dataset. These results emphasize the need for a more personalized approach to loneliness detection.

However, there were a few limitations with that work, including a small sample size, reliance on a single dataset, and a more static approach for detecting loneliness. These limitations highlighted the need for a larger dataset and a refined methodology to accommodate evolving behavioural patterns within the data better to improve the robustness of our approach. In this chapter, we aim to address these limitations and improve our methodology by using the GLOBEM dataset. The novelty of our work lies in adding incremental clustering with incremental classification on a weekly basis. This approach would help our models to evolve dynamically as new data becomes available each week, which will refine behavioural groups and loneliness detection over time. This adaptability is important for accurately capturing the fluid nature of student experiences and responding to loneliness early. We have also performed a temporal group analysis.

Our approach includes two primary components: incremental clustering and classification. The clustering part aims to identify subgroups of students with similar behavioural patterns. This process adds new weekly data in batches that refine the behavioural groups over time. The incremental classification component detects lonely individuals within these evolving student groups on a weekly basis. We maintain a main generic classification model trained on all available data, which is regularly updated on a weekly basis. Additionally, when new distinct behavioural patterns emerge in the form of a cluster, we create a specialized classification model for that specific group.

5.2 Methodology

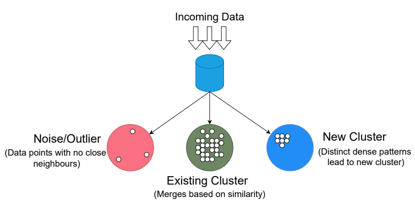
5.2.1 Evolving Student Behavioural Profiles via Incremental Clustering

We used the GLOBEM dataset for this study. To find behavioural groups of students on a weekly basis we used the incremental clustering approach with IDBSCAN as we did in Chapter 4. Unlike traditional clustering, which analyzes a static dataset, incremental clustering continuously adapts to new data without retraining the entire model. The incremental clustering process (illustrated in Figure 5.1a) begins with evaluating the new data against the existing clusters. It assesses whether the data closely aligns with the established behavioural pattern of a group (within the specified distance threshold eps) and if the group's density (number of points within eps) meets the minimum requirement (minPts). If the data fits with some existing group, it gets merged in it while increasing the cluster size. In case new data doesn't match with some existing cluster, it may indicate some new emerging behavioural trends or temporary fluctuations within the student population. When multiple data points show a distinct, consistent pattern, they form a new cluster. To determine the optimal values for eps (epsilon) and MinPts (minimum points), we used GridSearchCV combined with cross-validation and silhouette score as the evaluation metric. A set of candidate parameter values was initially selected based on experiments, as we did in Chapter 4 as well. The grid search explored these values for the eps and MinPts by evaluating each combination using the silhouette score as the performance metric. The best parameter values were chosen that maximized the average silhouette score across the cross-validation folds.

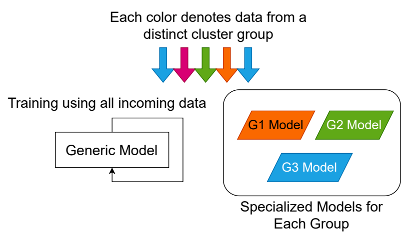
5.2.2 Dynamic Loneliness Detection through Incremental Classification

In our analysis of the StudentLife dataset in Chapter 4, we found the dynamic and evolving nature of students' behaviours. As discussed in Chapter 4, we implemented incremental clustering on a weekly basis to effectively monitor these behavioural changes in students and identified behavioural clusters that showed shifts in students' behaviours over time. Although we could track the formation and evolution of behavioural groups by using this approach but there was no timely loneliness detection within these groups. We were training classification models at the end of the study after week 10. To address this limitation, we are using the concept of incremental classification in this chapter, in which we can detect loneliness dynamically on an early basis based on the changes in students' behaviours over time. This approach can lead to a responsive detection that would detect loneliness in a timely and dynamic way.

Starting from the initial stage, we create a generic classification model trained on the initial data from week 1. This generic model is then updated incrementally with each week's incoming data from all students that allows it to adapt to general trends and evolving behaviours over time. As we presented in Chapter 4, a single generic model may struggle to capture the unique behavioural patterns associated with loneliness. To address this, we develop 'specialized' classification models tailored to each group identified through clustering which were trained exclusively on data from students within that subgroup. Fig.5.1b shows the process of refining the data that leads to the generic model and then to models G1, G2, and G3 which are each specifically made for a different group.



(a) Visualization of Dynamic Data Clustering: Segregation into Noise, Existing, and Emergent Clusters



(b) Personalized Loneliness Detection Using Behavioural Groups

Figure 5.1: Clustering and Personalized Classification

In our study, we used incremental classification for four ML algorithms: logistic regression, support vector machines (SVM), random forest, and XGBoost. For logistic regression and SVM, we used the Stochastic Gradient Descent (SGDClassifier) from scikit-learn due to its computational efficiency and it supports partial fitting for incremental learning [Dat24, WS16]. The SGDClassifier also includes L2 regularization by default which helps prevent overfitting by penalizing large coefficients. For logistic regression, we used log loss, while for SVM, we used the hinge loss function. For random forest, we used an Adaptive Random Forest (ARF) classifier from the river library [Dat22]. ARF is designed for incremental learning and incorporates drift detection mechanisms to identify and respond to changes in data patterns. For XGBoost, we used an incremental learning approach where the model was initially trained on the initial data, and then updated weekly with new data by using the train function with the previous model passed through the *xgbmodel* parameter [Vic23].

We implemented a weekly iterative approach to evaluate and refine our models over time. Each week starts with incremental clustering on the week's data to identify and update behavioural groups within the student population. These groups then lead to the training and updating of two types of classification models. First one is a single generic model trained on data from all students in that week which is trained to capture overall trends in data. The second are specialized models trained exclusively on data from students within each identified behavioural group. We applied the same data processing steps as we performed in Chapter 4. This included feature scaling, principal component analysis (PCA) for dimensionality reduction, handling of missing values and feature selection for classification models.

We start by stratified splitting the data into 95% for training/testing and 5% for validation. For the first six weeks, this validation set was accumulated for later hyperparameter tuning. In Week 1, we established initial models through stratified nested cross-validation with hyperparameter tuning using GridSearchCV. To address class imbalance, we used SMOTE within each cross-validation fold. For group-specific models, this process was repeated on the data specific to each cluster.

From Weeks 2 to 5, we continued incremental clustering on the data each week to update existing groups. Similarly, both generic model and group-specific models were updated incrementally with the new week's data. In parallel, the group-specific models were trained for any newly identified clusters. We evaluated generic models using 10-fold while group-specific models using 5-fold cross-validation. As we explained in Chapter 4, we have also not applied LOSO CV in this chapter due to the small sample sizes in each week's groups.

In Week 6, due to observed performance drops, we re-tuned the hyperparameters for all models using the accumulated validation data (30%) and nested cross-validation. This adjustment ensured that the new hyperparameters generalized well across weeks.

From Weeks 7 to 10, we continued the weekly update and evaluation process for all models. In the final week, we conducted a comprehensive analysis of the performance trends across the entire 10-week study period which we will present in the results section. We evaluated models using the average performance metrics from the cross-validation process. The key performance metrics included accuracy, precision, sensitivity, and F1-score.

5.3 Results

5.3.1 Temporal Evolution of Student Groups and Behavioural Profiles

Fig. 5.2 shows the dynamic shifts in student behavioural groups over a 10-week study period. This figure shows the temporal evolution of these groups with each group denoted by a unique identifier (G1, G2, G3, G4). The y-axis represents the number of students in each group, while the x-axis represents each week. To track group membership changes over time, each student was assigned to the cluster with the highest count of their data points within a week, which classified them into the group where their behaviour most frequently aligned.

Over the ten weeks, students switched between groups as their behavioural patterns changed. This means that group membership was not fixed, and students could shift from one group to another, which reflects their evolving behaviours. For example, a student who showed highly social and active behaviours in the early weeks may belong to G2, but if their behaviour becomes more isolated and less active, they may move to G3 or G4 in later weeks. These transitions capture the fluid nature of student behaviour, showing how their engagement levels and social interactions vary over time, which is essential for understanding how loneliness manifests differently across individuals based on their own subjective experiences.

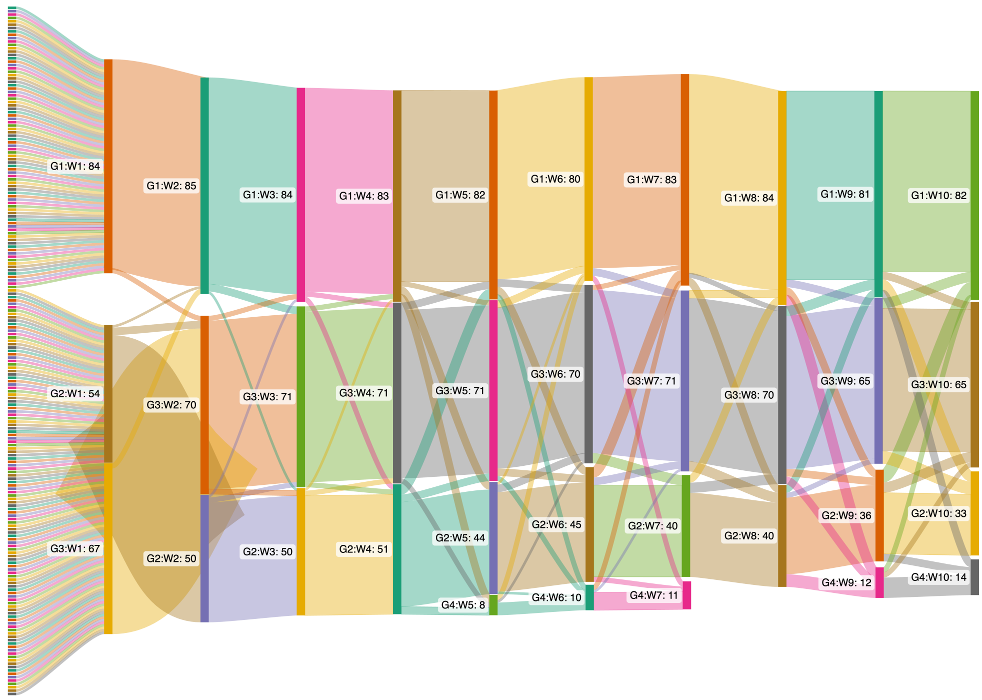


Figure 5.2: Visualization of Students Temporal Behavioural Group Dynamics over a 10-Week Period (G1:W1:84 denotes Group 1 in Week 1 with 84 students)

5.3.1.1 Behavioural Profiles

We categorized the weekly behavioural patterns of students within each group using globally normalized feature values using min-max scaling across the entire 10-week period. For each feature, values below the 25th percentile were categorized as 'Low' values between the 25th and 75th percentiles as 'Moderate' and values above the 75th percentile as 'High'. This approach ensures consistent interpretation of categories across all weeks. This leads to the identification of potential trends in loneliness scores associated with different behavioural profiles. Table 5.1 summarizes these observations and provides detailed information on the behavioural tendencies and associated loneliness scores for each group. Behavioural patterns are summarized for the week intervals which shows the overall dominant patterns during those weeks. However, in the cases where some variations occurred, we have added symbols with the week number to show those variations. For example, the entry 'High calls & SMS (=W2)' for Group 1 during weeks 1-4 indicates that call and SMS activity was generally high during this period, with a moderate level specifically in week 2.

Table 5.1: Summary of Behavioural Patterns and Associated Loneliness Scores by Group Over 10 Weeks (($\uparrow$ = High, = = Moderate, $\downarrow$ = Low))

Week Interval	**Group**	**Behavioural Patterns**	**Loneliness Score Range**
Weeks 1-4	G1	High physical activity, High calls & SMS(=W2), High unique Bluetooth encounters, Moderate location changes, Moderate phone usage duration($\uparrow$W1), High sleep duration(=W3)	10-18
	G2	High physical activity(=W1), Moderate calls & SMS, Moderate unique Bluetooth encounters($\downarrow$W2), Moderate phone usage($\downarrow$W4), Moderate sleep duration($\downarrow$W3)	15-26
	G3	Low physical activity, Low calls & SMS, Low unique Bluetooth encounters, Moderate location changes($\downarrow$W1), Moderate phone usage($\uparrow$W3), Low sleep duration	16-30
Weeks 5-7	G1	High physical Activity(=W7), Moderate calls & SMS, High unique Bluetooth encounters(=W5), Moderate location changes, Moderate phone usage($\downarrow$W6), High sleep duration(=W7)	10-22
	G2	Low physical activity, Low calls & SMS, Moderate unique Bluetooth encounters, Moderate location changes, Moderate phone usage($\uparrow$W7), Low sleep duration(=W5)	15-32
	G3	Low physical activity, Low calls, Moderate unique Bluetooth encounters, Low location changes(=W7), High phone usage, Moderate Sleep duration($\uparrow$W5)	18-36
	G4	Moderate physical activity($\downarrow$W6), Moderate calls & SMS, High unique Bluetooth encounters, Moderate location changes($\downarrow$W7), Moderate phone usage, Moderate sleep duration($\downarrow$W6)	14-28

Continued on next page...

Table continued...

Week Interval	**Group**	**Behavioural Patterns**	**Loneliness Score Range**
Week 8	G1	Moderate physical activity, High calls & SMS, High unique Bluetooth encounters, Moderate location changes, Moderate phone usage, Moderate sleep duration	10-20
	G2	High physical activity, Moderate calls & SMS, Moderate unique Bluetooth encounters, Moderate phone usage, Moderate sleep duration	15-25
	G3	Low physical activity, Moderate calls & SMS, Low unique Bluetooth encounters, Moderate location changes, Moderate phone usage, Low sleep duration	18-36
Weeks 9-10	G1	High physical activity, High calls & SMS, High unique Bluetooth encounters, Moderate location changes, Moderate phone usage(↓W9), High sleep duration	10-20
	G2	Moderate physical activity(↓W9), Moderate calls & SMS, Moderate unique Bluetooth encounters, Moderate phone usage, Moderate sleep duration	16-32
	G3	Low physical activity(=W10), Moderate calls & SMS, Low unique Bluetooth encounters, Moderate location changes, High phone usage(=W10), Low sleep duration	18-36
	G4	Low physical activity, Moderate calls & SMS, Moderate unique Bluetooth encounters(↓W9), Moderate location changes, Moderate phone usage, Moderate sleep duration	14-35

Initial Phase (1-4)

During the initial weeks, three groups emerged. G1 students showed overall high physical activity which is calculated by steps count, high number of calls/SMS, and diverse social interactions (Bluetooth encounters), which shows good engagement with a balanced, active schedule and high sleep. Their loneliness scores ranged from 10 to 18, which shows that their active lifestyle may correlate with lower feelings of loneliness. Similarly, G2 was physically active but with moderate social interaction and phone use. This points to an active lifestyle, but perhaps less socially intensive than G1. In contrast, G3 showed the lowest physical activity, limited calls/texts, and fewer Bluetooth encounters, alongside heavy phone use and shorter sleep. Their loneliness scores ranged from 16 to 30, which is higher in comparison with other 2 groups.

Middle Phase (Weeks 5-7)

By the middle phase of the study, behavioural patterns changed a little bit. G1 maintained their patterns which shows stability within this group. However, G2 showed a decline in activity, social interaction, and sleep. This group showed a shift in score, with increased loneliness scores (15-32). G3 remained largely unchanged, with persistently low activity, short sleep, and heavy phone usage. The emergence of G4 introduced further subtlety; this group showed moderate activity, moderate communication patterns, and moderate sleep.

Final Phase (8-10)

In the final weeks, G1 and G2 showed moderate to high physical activity and communication patterns and moderate sleep. However, G3 remained low physically active with moderate phone usage, while G4's physical activity levels decreased.

G4 disappeared at week 8 but reappeared at week 9. To determine if the G4 that emerged in Week 9 shared similar characteristics with the earlier G4 from Weeks 5-7, we conducted a feature average analysis. We calculated the mean and standard deviation for each feature in both periods. Our analysis found that the majority of feature averages from Week 9 fell within one standard deviation of their corresponding averages from Weeks 5-7.

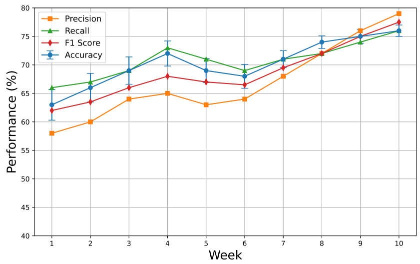
5.3.2 Loneliness Detection Using Generic Models

Table 5.2 presents the performance of four generic classification models at the end of the study, which have been trained on a weekly incremental basis. The table shows the performance of XGBoost, Random Forest, SVM, and Logistic Regression models for the lonely class. As the table shows, the XGBoost model achieved the highest accuracy (76.52%) followed by the Random Forest with 74.64% and then SVM with 63.50% and then the Logistic Regression model with an accuracy of 57.32%.

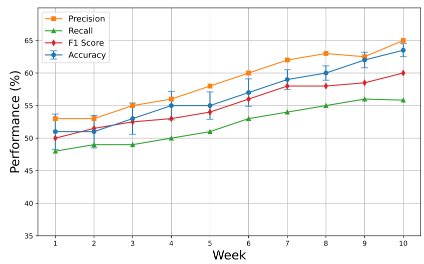
Table 5.2: Performance of Generic Models for Loneliness Detection

Model	**Accuracy**	**Precision**	**sensitivity**	**F1 Score**
XGBoost	76.52%	79.00%	76.83%	77.50%
Random Forest	74.64%	71.50%	66.72%	69.58%
SVM	63.50%	65.00%	55.50%	60.12%
Logistic Regression	57.32%	43.50%	54.65%	48.53%

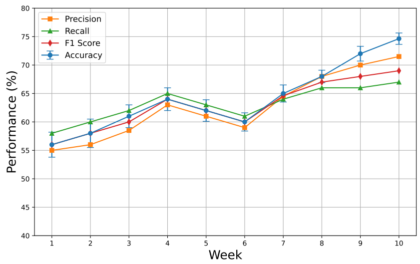
Figure 5.3 shows a more nuanced perspective on how these models performed over the 10-week study period. We can see that each model showed a trajectory of performance in the prediction of loneliness. Although there were initial fluctuations but the final week's results show that the models capabilities improved gradually. The XGBoost model showed gradual improvement, starting at 63% accuracy in week 1 and reaching 76.5% by week 10, with corresponding increases in precision, sensitivity, and F1 score to 79%, 76.83%, and 77.50%, respectively. The model performance displayed a minor dip around weeks 5 and 6, which was subsequently improved. Random Forest accuracy improved from 59% to 74.64%, with precision at 71.5% and sensitivity at 66.72%. The F1 score for this model was 69.58%. Similar to other models, Random Forest shows weekly variations specifically around weeks 5 and 6. The SVM model's accuracy improved from an initial 52% to 63.5%. The logistic regression model shows a performance increase over a 10-week period, with accuracy starting below 40% and rising to over 57.32%.



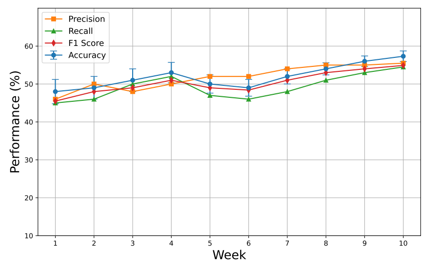
(a) XGBoost Model



(b) Support Vector Machine Model



(c) Random Forest Model



(d) Logistic Regression Model

Figure 5.3: Temporal Performance of Generic Classification Models for Loneliness Detection Across a 10-Week Period

5.3.3 Loneliness Detection Using Group Based Models

In this chapter, we trained loneliness classification models in a weekly incremental manner meaning that models were updated and trained each week based on the data from the dynamically evolving behavioural groups. This approach provides a more real-time and adaptive detection of loneliness compared to the static and retrospective classification approach used in Chapter 4. By incorporating weekly behavioural data, the models can better reflect the current state of a student's behaviour that allows for timely detection of loneliness as students shift between groups. This weekly training ensures that the model adapts to emerging trends in behaviour and can capture short-term changes that may indicate increasing or decreasing loneliness. Instead of relying on cumulative behaviour across the entire study as in Chapter 4, the weekly approach ensures that predictions are responsive to the latest behavioural patterns which provides a more precise and personalized detection approach. This incremental method allows the model to dynamically adjust to new behaviours as they emerge which improves its ability to detect loneliness early and more accurately compared to a static model trained only once at the end of the study.

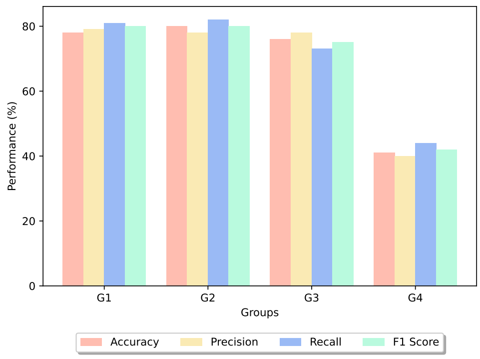
The final results from the group-based models for loneliness detection at week 10 are presented in Table 5.3 and Figure 5.4. XGBoost performed best with the highest accuracy in Groups 1 and 2, with scores of 78.64% and 80.15%, respectively. Its precision, sensitivity, and F1 scores were also high. Group 3 show a slight drop in accuracy for XGBoost to 76.55%. However, in Group 4, the performance was poor, with accuracy falling to 41.38%. The Random Forest algorithm showed almost similar accuracy as of XGBoost for G1 (78.35%) but all other performance metrics were low. For Group 4, the Random Forest algorithm outperformed XGBoost, achieving a higher accuracy of 62.42% and an F1 score of 64%. The SVM model's performance was relatively lower, with the highest accuracy of 66.43% for G1. The logistic regression model showed the lowest performance across all groups, with accuracy scores ranging from 42.48% to 64.29%.

Adding error bars to the accuracy helped show that some of the drops in performance during the middle weeks may not be significant since the performance ranges overlap with nearby weeks. This means the model's performance is likely improving as more data is added, but there are some ups and downs due to the natural variability in the data and the way the model is trained. These changes are more noticeable in the earlier weeks, where there is less data available for training.

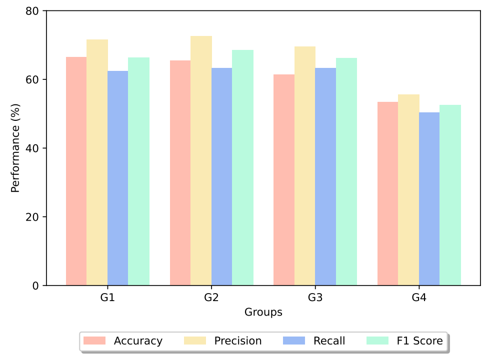
Table 5.3: Performance of Group Based Models Classifying Loneliness at Week 10

Algo	G1				G2				G3				G4			
	Acc.	Prec.	Sens.	F1.	Acc.	Prec.	Sens.	F1.	Acc.	Prec.	Sens.	F1.	Acc.	Prec.	Sens.	F1.
XGBoost	78.64	79.84	81.5	80.5	80.15	78.63	82.44	80.74	76.55	78.34	73.74	75.97	41.38	39.66	44.38	42.10
RF	78.35	77.74	74.44	75.73	79.00	80.45	78.55	79.34	75.53	78.43	80.14	79.55	62.42	63.40	65.33	64.00
SVM	66.43	71.52	62.45	66.38	65.49	72.55	63.32	68.43	61.31	69.52	63.33	66.18	53.41	55.53	50.29	52.52
LR	64.29	54.32	63.71	58.45	58.29	55.62	58.41	56.29	52.00	44.42	55.29	48.73	42.48	39.84	44.23	41.96

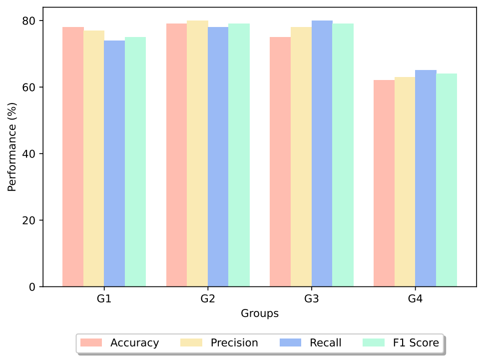
Algo:Binary Classification Algorithm; G1: Group 1; G2:Group 2; G3:Group 3
G4:Group 4; Acc:Accuracy(%); Prec:Precision(%); Sens:sensitivity(%) F1:F1 Score(%)
RF:Random Forest; SVM:Support Vector Machine; LR:Logistic Regression



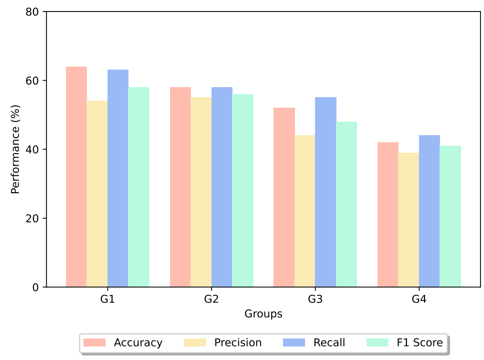
(a) XGBoost Model



(b) Support Vector Machine Model



(c) Random Forest Model



(d) Logistic Regression Model

Figure 5.4: Performance of the Group based Classification Models for Loneliness Detection at Week 10

5.3.4 Aggregated Performance of Personalized Models

To assess the overall effectiveness of group-based approach for loneliness detection compared to the generic models, we calculated the aggregated performance metrics for the personalized models as we did in chapter 4 as well. We used a weighted average based on the number of data points in each group to ensure that the aggregated metrics accurately reflect the contributions of each group. Table 5.4 presents the performance metrics for aggregated personalised models.

Table 5.4: Aggregated Performance Metrics of Personalized Models Across All Groups

Model	Accuracy,%	Precision,%	Sensitivity,%	F1 Score,%
XGBoost	76.10%	76.25%	77.79%	77.04%
RF	76.92%	77.74%	76.09%	76.72%
SVM	64.28%	70.38%	62.04%	65.95%
LR	58.62%	52.05%	59.08%	55.34%

The aggregated personalized XGBoost model had a slightly lower accuracy compared to the generic model by 0.42% (76.10% vs. 76.52%). Precision decreased by 2.75% (76.25% vs. 79.00%), but sensitivity increased slightly (77.79% vs. 76.83%). The F1 score showed a minor decrease (77.04% vs. 77.50%). The aggregated personalized RF model improved accuracy by 2.28% (76.92% vs. 74.64%), precision by 6.24% (77.74% vs. 71.50%), sensitivity by 9.37% (76.09% vs. 66.72%), and F1 score by 7.14% (76.72% vs. 69.58%). The personalized SVM and Logistic Regression models both showed minor improvements in accuracy, precision, sensitivity, and F1 score compared to the generic models.

5.4 Discussion

5.4.1 Temporal Dynamics of Loneliness

While Chapter 4 presented about the dynamic nature of student behaviour and loneliness, the larger dataset in this chapter helped us to investigate a more nuanced and comprehensive understanding of these temporal patterns and how they relate to loneliness levels. Over the study period, we have seen shifts in group composition and behaviours of students in each group. This evolution highlights the importance of a dynamic approach to monitor loneliness, which may provide indicators of the temporal aspects of student lifestyles and their mental well-being as we discussed in Chapter 4 as well. Although the temporal analysis in Chapter 4 suggested the possibility of external factors affecting changes in behaviour, the longer time period analyzed in this chapter provided a more detailed investigation of these patterns. We observed periodic fluctuations in group composition and behaviour that might align with typical academic cycles. These changes could potentially be associated with periods of increased academic stress, such as those that might occur during midterms or final exams. However, further studies are needed with contextual data to explore the causal relationships between these external factors and behavioural changes to understand why these patterns emerge definitively. Monitoring these changing patterns can helps us understand how stress, social interactions, and personal circumstances affect student behaviour.

Group 1 maintained consistently high physical activity and social interactions with some fluctuations throughout the study which might align with their relatively stable and low loneliness scores (10-22). This shows that a balanced lifestyle potentially contributes to lowering feelings of loneliness. However, Group 2 shows fluctuations in behaviour with decreased activity and social interactions during weeks 5-7, with increased phone usage. This pattern may indicate that academic or personal stressors negatively impact social engagement and well-being, although further research is needed to find the actual reasons. The emergence of Group 4 in week 5 showed moderate behaviours and varied loneliness scores (14-35). This shows the diverse ways students adapt to changing academic and social demands. Group 3 consistently showed low physical activity, high phone usage, and higher loneliness scores (16-36), which aligns with previous research linking reduced physical engagement and increased technology use to social withdrawal and loneliness [HTC09, NNC18]. The behavioural shifts in weeks 9-10, specially for Groups 2 and 3 could be due to end-of-semester pressures, which might have impacted social behaviours and loneliness levels.

5.4.2 Effectiveness of Personalized Models

The results from this chapter further highlight the effectiveness of personalized models in loneliness detection among students. The performance of group specific models over generic models, previously observed in the smaller dataset has again been validated in this chapter. Personalized models excel in performance due to their ability to detect nuanced behavioural patterns that are often masked in generic models trained on the whole dataset. Tree-based models outperformed other algorithms, likely due to the fact that they work better to handle non-linear relationships and complex interactions between features. Their ensemble nature helps them to capture a wider range of patterns in the data that potentially lead to improved predictive performance which we can see in the results. The underperformance of all models in Group 4 shows some difficulties in predicting loneliness for this group. This group might have more complex behaviour patterns that are harder for the model to learn effectively related to lonely class. There could also be data quality issues specific to this group which results in poor model performance. Despite the marginally lower performance of the aggregated personalized XGBoost model compared to its generic counterpart, the overall findings support our hypothesis that personalized models are more effective, especially when we analyze them individually for each group. However, it is important to note that the improved results, compared to the generalized approach, may be influenced by the use of traditional 10-fold cross-validation rather than leave-one-subject-out cross-validation (LOSO), which could have provided a more robust evaluation.

The choice between generic and personalized models involves a trade-off. Generic models might be less precise for individual groups, but they can be used in a broader applicability. They can be more suitable when resources are limited or where a rapid, general assessment is needed. On the other hand, personalized models work well in situations where targeted interventions are required and a deeper understanding of individual differences is important. For example, in a university setting, a generic model could be used to identify students at risk of loneliness broadly, while group-specific models could be used for designing tailored interventions for specific student populations. We will be discussing this tradeoff between personalisation and generalisation in detail in Chapter 8 (section 8.2.1).

5.4.3 Comparison with Chapter 4 Findings

The comparison of loneliness detection models between Chapter 4 (StudentLife dataset, 41 students) and Chapter 5 (GLOBEM dataset, 205 students) provides important observations about the challenges and complexities of scaling up passive sensing based loneliness detection systems.

We observe a significant decrease in performance across all metrics for generic models when moving from the smaller StudentLife dataset to the larger GLOBEM dataset. For instance, the XGBoost model's F1 score dropped from 88.50% to 77.50%, with similar declines in accuracy, precision, and sensitivity scores. This trend is consistent across all algorithms with more significant decreases for simpler models like SVM and logistic regression. This decline indicates that increasing sample size introduces more complexity in loneliness detection. Although this increase in sample size from 41 to 205 students provided more data for models to train but this may have introduced greater variability in behavioural patterns associated with loneliness. This increased complexity could have made it more challenging for models to learn the specific patterns associated with 'lonely' class and generalize effectively. Secondly, the incremental training approach used in this chapter, where models were updated weekly with new data could also have contributed to reduced performance. Although this approach is more reflective of real-world scenarios but this may struggle to capture and remember consistent patterns related to loneliness as effectively as training models with all data at once.

The group-specific models also show a similar trend of decreased performance when moving from the StudentLife to the GLOBEM dataset. However, the group-specific models consistently outperform their generic counterparts in both datasets which highlights the importance of personalized approaches in loneliness detection. The observed decline in performance with the larger GLOBEM dataset aligns with findings from a recent depression detection study using passive sensing data [ASM+24]. Similar to our findings, this study found that ML models have lower accuracy in larger, more diverse samples. This is due to the varying meanings of behavioural indicators for depression across different demographic and socioeconomic subgroups. This phenomenon likely extends to loneliness detection as well since the underlying mechanism is the same. Just like they found different associations between phone usage and mobility patterns with depression across age and socioeconomic groups, similar behavioural indicators of loneliness may vary in their predictive power across different student subpopulations. Our group-specific models' superior performance compared to generic models across both datasets supports this theory. It indicates that loneliness just like depression manifests differently across various behavioural subgroups. The more apparent performance decline in generic models when moving to the larger GLOBEM dataset further supports this idea, as a larger sample size likely introduced greater variability in how behavioural patterns relate to loneliness. We will discuss this model generalizability aspect in more detail in Chapter 8, 'Conclusion'.

5.4.4 Implications for Interventions and Personalized Support

Together, the findings from Chapter 4 and Chapter 5 present practical applications for personalized loneliness detection among students. The dynamic nature of student behaviour and loneliness presented in our study highlighted the importance of continuous monitoring and adaptive interventions. As students navigate through different academic and social challenges, their risk of loneliness and its behavioural manifestations may change. By continuously tracking these changes, universities can ensure that their support services remain relevant and effective throughout the academic year. Moreover, generic interventions often lack the specificity needed to be truly effective. Personalized detection can lead to the development of targeted support systems tailored to the unique behavioural patterns of different subgroups in a broader population. For example, interventions for groups with high phone usage and low physical activity could focus on encouraging physical engagement and face-to-face interactions. This approach can improve early interventions by continually monitoring behavioural patterns and detecting shifts associated with loneliness, which would lead to timely and personalized support for students suffering from loneliness. Another important aspect to highlight is while smartphones and wearables provide passive sensing data for monitoring and detecting loneliness, they may also contribute to increased loneliness, specifically among younger populations [EDKN+16]. This dual implication highlights the importance of balancing the benefits of using these technologies for data collection with their potential negative impacts.

5.5 Conclusions and Limitations

This chapter has expanded upon the foundations laid in Chapter 4, addressing its limitations and further validating the efficacy of personalized, dynamic approaches to loneliness detection among students. Our main findings are:

- The usage of incremental clustering and classification has proven to be a key improvement in our approach. This approach allowed models to adapt dynamically to evolving behavioural patterns that provide a timely representation of student experiences. This adaptability is important in capturing the fluid nature of loneliness and its behavioural indicators which ultimately will lead to early detection and intervention.
- The results of this study further reinforce the findings from Chapter 4 which emphasizes the effectiveness of personalized models over generic ones in detecting loneliness.
- The temporal weekly analysis of students' behaviour has shown the dynamic nature of students' lives and its relation with their loneliness levels. The group fluctuations and behavioural patterns highlight the need for continuous monitoring and adaptive intervention strategies.
- We observed a decrease in model performances when coming from the smaller StudentLife dataset to the larger GLOBEM dataset. This shows that increasing sample size and diversity added greater complexity to loneliness detection, emphasizing the need for further research.

In Chapter 8, we will discuss possible future work related to loneliness detection systems. Our discussion there will partially build upon the following limitations that we identify in the work of this chapter.

- In this chapter, we performed weekly binary classification of loneliness based on the students' behavioural data. However, it is important to clarify that the ground truth used for these classifications, derived from the UCLA Scale, was collected at the end of the 10-week study period. This means that the UCLA scores provide a cumulative measure of each student's loneliness at the end of the study, rather than reflecting their loneliness on a weekly basis. The weekly classifications rely on using this post-study score as a consistent label across all weeks. While this approach is common when longitudinal ground truth data is unavailable, it introduces a limitation about capturing precise weekly fluctuations in loneliness. Future studies should focus on collecting loneliness scale data more frequently throughout the study period, to better track temporal changes and capture a more precise measurement of loneliness over time.

- As we highlighted in Chapter 4 as well, the dataset we are using is limited to a specific student population from a single university over a fixed period. The homogeneity of this sample may not fully capture the diverse range of experiences and behavioural manifestations of loneliness present in the broader population.
- Our approach worked well, but there is still room for improvement, especially in predicting loneliness within groups like G3 and G4. To address this, future work needs to explore the potential reasons for reduced performance and may use advanced methods to improve performance. Also there is a need to test LOSO CV for more robust evaluation. Moreover, ensuring such detection models are fair and robust is important. This means they should work equally well for all subgroups without any bias specific to certain populations' behaviours or demographics. More contextual data related to populations will need to be incorporated to train these models effectively. Fairness in these models will make them more reliable and also help build trust in such systems' ability to support every student effectively.

Chapter 6

Differentiating and Predicting Social and Emotional Loneliness in Students Using Passive Sensing: An Exploratory Study

6.1 Overview

Previous chapters have explored multiple aspects related to loneliness, including its behavioural manifestations and detection techniques. While loneliness is often viewed as a single emotional experience, it is far more complex [Wei75]. In this chapter, we will explore its multifaceted nature, mainly by examining the distinction and overlap between social and emotional loneliness. The social dimension of loneliness refers to the subjective experience of feeling socially disconnected and lacking a sense of belonging [RCRY84]. This is commonly seen when individuals have fewer personal relationships or a smaller social circle than they would want. On the other hand, the emotional element emphasizes the lack of a deep, personal level connection with others, which could be deeply experienced even in the physical presence of others [Vau88]. The UCLA scale has been used by researchers to investigate the interconnected aspects of social and emotional loneliness [HBC05, CHE+06]. Studies have indicated that experiences of social and emotional loneliness do not always coincide. For instance, a person might have a robust peer network yet miss having a close friend, which indicates they feel emotional loneliness without necessarily feeling socially isolated

[MVVdNG17, QM02]. This distinction is critical for a nuanced understanding of loneliness, as social and emotional loneliness may arise from different circumstances, have distinct physiological and psychological effects, and therefore may require different intervention strategies [LGB+11]. Specifically, emotional loneliness has been linked with emotional challenges, including negative self-views, introverted behaviour, and distorted views of relationships [Rok19]. While prior research has identified these distinct types, to the best of our knowledge, no study has investigated how they manifest in behavioural patterns derived from passive sensing data within the student population specifically.

This chapter aims to investigate social and emotional loneliness using passive sensing data for a student population. Our goal is to determine the subtle differences in data from individuals who are socially and emotionally lonely, evaluate the predictive power of behavioural features in categorizing these loneliness types, and find the behavioural features that are most important for differentiating between these two types of loneliness. The key research questions are:

1. Can behavioural features extracted through passive sensing data differentiate between socially and emotionally lonely students?
2. What is the power of behavioural features to classify social and emotional loneliness?
3. What behavioural features are most important for predictive models in classifying loneliness and its types?

6.2 Methods

6.2.1 Categorizing Social and Emotional Loneliness

We have used the GLOBEM dataset for this study. In the dataset, loneliness was assessed using the revised 10-item UCLA Loneliness Scale [KCMG88]. The revised 10-item UCLA Loneliness Scale uses a 4-point response scale for each item, where 1 represents 'never', 2 'rarely', 3 'sometimes', and 4 'often'. With 10 items in total, the minimum possible score is 10 (if a participant responds with 1 for all 10 items), and the maximum possible score is 40 (if a participant responds with 4 for all 10 items). Therefore, the overall score range for this 10-item questionnaire spans from 10 to 40. For our research purposes, we divided the items of scale into two distinct categories of social and emotional loneliness based on the criteria proposed in [MQLM22, BPR+08]. Items in the scale that

Table 6.1: Division of 10-item UCLA scale into Emotional and Social Loneliness, 'R' indicates reverse scoring

Type	Questions
Emotional Loneliness	1. How often do you feel that no one really knows you well? 2. How often do you feel close to other people? (R) 3. How often do you feel that there are people who really understand you? (R) 4. How often do you feel that there are people you can turn to? (R) 5. How often do you feel that people are around you but not with you?
Social Loneliness	1. How often do you feel that you have a lot in common with the people around you? (R) 2. How often do you feel that you feel left out? 3. How often do you feel isolated from others? 4. How often do you feel that there are people you can talk to? (R) 5. How often do you feel that you lack companionship?

are related to feeling isolated, feeling left out, lack of companionship, and social interactions are classified under the social loneliness category, while the items related to emotional disconnect, like not feeling close to others, and no one truly understanding me, are put under the emotional loneliness category. Each category consisted of 5 items from the original scale presented in Table 6.1. Some items in the scale, denoted by (R), are reverse-scored. This means that for these items, the scoring is inverted: a response of 1 is scored as 4, 2 as 3, 3 as 2, and 4 as 1. This reverse scoring ensures that higher scores consistently indicate greater levels of loneliness across all items.

There is no single universally accepted threshold in the literature for determining loneliness cutoff scores for dividing into low or high loneliness categories. Many studies have proposed their own cutoff scores like one by Cacioppo et al. [CP08]. In our study, participants were scored on two dimensions: social loneliness (referred to as social_score) and emotional loneliness (referred to as emotional_score), each ranging from 5 to 20. A cumulative score of 10 is achieved if a participant selected 'rarely' for all 5 questions, which represents occasional feelings of social or emotional loneliness. Hence, we chose 10 as our cutoff score for each subscale. To categorize loneliness levels, we used the following approach:

- Participants with a social_score of above 10 and an emotional_score of 10 and below were labeled as 'socially lonely'.
- Participants with an emotional_score of above 10 and social_score of 10 and below were labeled as 'emotionally lonely'.
- Participants scoring above 10 on both scales were considered 'both socially and emotionally lonely'.
- Finally, those scoring 10 or below on both scales were categorized as 'not lonely'.

We first applied some basic statistics to analyze the different aspects of loneliness among the students in the dataset. This involved calculating the mean, median, first quartile (Q1), third quartile (Q3), and standard deviation (SD) for the four distinct categories: 'socially lonely', 'emotionally lonely', 'both socially and emotionally lonely', and 'not lonely'. This provided an initial overview of the overall distribution and central tendency of loneliness within the dataset.

6.2.2 Differentiating Social and Emotional Loneliness

We conducted a statistical analysis using behavioural features to address our first research question about differentiating between social and emotional loneliness. Before applying statistical analysis, we used Mutual Information (MI) for feature selection to identify the most relevant behavioural features for differentiating between social and emotional loneliness. We selected features whose cumulative MI reached a threshold of 95% of the total MI across all features.

We then checked the normality of the distribution of behavioural features for both socially and emotionally lonely student groups. This was conducted using the Shapiro-Wilk test [Roy83], which is a robust method for testing normality in data and it is particularly effective for small sample sizes like ours (24 participants in the socially lonely group and 19 in the emotionally lonely group). The normality check is important because different statistical tests are suitable for different types of data distributions. Parametric tests, such as t-tests or ANOVA, assume that the data follows a normal distribution. If this assumption is violated, non-parametric tests like the Mann-Whitney U test, are more appropriate as they do not rely on the normality assumption. By using the Shapiro-Wilk test, we ensured the validity and reliability of our subsequent statistical analyses in identifying differences between socially and emotionally lonely groups.

Given that the Shapiro-Wilk test indicated a non-normal distribution for majority features, we used a two-sided Mann–Whitney U test to compare the feature distributions between the socially and emotionally lonely groups. This non-parametric test was selected for its ability to compare differences between two independent samples without the assumption of normal distribution. The goal was to determine if there were statistically significant differences in the feature profiles between socially lonely and emotionally lonely groups. The null hypothesis for this test was that there was no difference in the mean of features between the two groups, while the alternative hypothesis stated that the mean of one or more features would differ.

To quantify the magnitude of the differences observed between the socially and emotionally lonely groups, we calculated effect sizes using the statistical technique called bootstrapping. Bootstrapping is a resampling technique that enhances the reliability of statistical inferences, which is particularly useful when working with smaller sample sizes. This method works by repeatedly sampling with replacement from the original dataset to create multiple simulated datasets. This helps for the estimation of sampling distributions and the calculation of robust statistics. In our case, a total of 10K bootstrap samples were used to estimate the distribution of the effect size and resampled features with replacement, which provides a more robust measure in the context of our small sample sizes. This involved resampling the observations (with replacement) 10K times to create multiple bootstrap samples. We selected Cohen's d as our effect size metric. Cohen's d is a standardized measure of the difference between two group means, that is expressed in units of standard deviations. This standardization is used for meaningful comparisons across different variables with varying scales which makes it suitable for our analysis of diverse features. Alongside the point estimate, a 95% confidence interval was computed from the bootstrap distribution to calculate the precision of the effect size.

6.2.3 Predictive Modeling for Loneliness Classification

We trained multiple ML multi-class classification algorithms to address our second research question regarding the power of behavioural features in classifying loneliness types. The dataset has instances labeled as 'socially lonely', 'emotionally lonely', 'both lonely', or 'not lonely'. As we explained in Chapter 3 as well, we selected four widely used machine learning algorithms for our classification task: Support Vector Machine (SVM), XGBoost, Random Forest, and K-Nearest Neighbors (KNN).

We used nested cross-validation for model evaluation and selection. In the outer loop, we used leave-one-subject-out cross-validation to assess generalization. Within each iteration of the outer loop, we used an inner loop of stratified three-fold cross-validation for hyperparameter tuning using GridSearchCV. Preprocessing steps were conducted within each inner fold before model training and evaluation, including feature scaling, handling missing values, feature selection using MI, and handling class imbalance with SMOTE. We used the macro-averaged F1 score as the evaluation metric for model selection in cross-validation.

For baseline comparisons, we considered the majority class, random weighted classifier, and decision tree models trained on the original, imbalanced dataset to provide a fair comparison. The majority class model always predicts the most frequent class in the dataset to serve as a naive baseline. The random weighted classifier makes predictions randomly. The decision tree model makes classifications based on a series of feature-based questions to serve as a middle-ground comparison between the simplest baselines and our more complex models. These baselines serve as reference points to evaluate the performance of more complex models. The primary metrics for assessing the performance of each model were accuracy, precision, sensitivity, and F1 score. These metrics were calculated for each loneliness category which provide a detailed view of the model's performance. Accuracy provided a global view of model performance, while precision, sensitivity, and the F1 score provided insights into the model's ability to predict each specific class.

6.2.4 Feature Importance Analysis

To address our third research question regarding determining the most important features used by the predictive models in distinguishing between the different loneliness types, we used the Shapley Additive exPlanations (SHAP) values for XGBoost and Random Forest models [LEC+20]. We chose these two models specifically because they provided the best results in our predictive classification task.

SHAP values measure the impact of each feature on the model's output for classification. In our multi-class context, SHAP computes values for each class separately to compute feature importance for each loneliness type. Once the classification models were trained, we used the SHAP library to compute the SHAP values for each feature across all data points. This process outputs a SHAP value for each feature for each prediction and each class, indicating the feature's impact on the model's output for classification. We averaged the absolute SHAP values for each feature within each class to analyze the importance of class-specific features. Averaging the absolute SHAP values helped us to see the overall influence of each feature on the model's predictions to provide a clear and robust measure of feature importance within each class. We then ranked the features based on their average absolute SHAP values within each class to identify the most important behavioural indicators for each loneliness category.

6.3 Results

6.3.1 Overview of Loneliness in Participants

For the specific dimensions of loneliness, the mean social loneliness score was 10.93, with a median of 11 and an interquartile range (Q1: 9, Q3: 13) with a standard deviation of 2.736. Emotional loneliness scores have a mean of 10.71, a median of 11, and the same interquartile range (Q1: 9, Q3: 13), but a slightly higher standard deviation of 2.905. These statistics include all students to capture the full range of social and emotional loneliness within the dataset.

When classifying participants based on their social and emotional loneliness scores, we found a bit larger subset experiencing social loneliness (11.71%, 24 out of 205) compared to emotional loneliness (9.27%, 19 out of 205). However, a significant portion (42.44%, 87 out of 205) reported feeling both socially and emotionally lonely which highlights the interconnectedness of these experiences. 36.59% (75 out of 205) were classified as feeling neither socially nor emotionally lonely.

6.3.2 Statistical Differences for Social and Emotional Loneliness

To determine whether behavioural features could differentiate between social and emotional loneliness, we conducted a two-sided Mann-Whitney U test on the non-normally distributed data (as confirmed by the Shapiro-Wilk test). Table 6.2 presents the statistically significant ($p < 0.05$) features only with mean differences and their effect sizes for socially and emotionally lonely groups. We interpret the magnitude of effect sizes (Cohen's d) using the commonly used approach, where values of 0.2, 0.5, and 0.8 are considered as thresholds for small, medium, and large effects, respectively [Coh13]. Below are the key results.

Table 6.2: Summary of Significant Features for SL and EL Groups Including Mean Differences and Effect Sizes.

Features	**SL Group**	**EL Group**	**MDiff**	**Effect Size (Cohen's d)**
		Location		
LogLocationVariance(evening)	2.301 (1.864, 2.875)	3.751 (3.284, 4.324)	-1.452	-0.715 (-0.964, -0.514)
NumberOfSignificantPlaces	1.504 (1.67, 2.53)	2.167 (1.675, 2.602)	-0.663	-0.327 (-0.529, -0.264)
NumberOfLocationTransitions	5.463 (4.174, 6.732)	7.374 (5.638, 9.532)	-1.911	-0.780 (-0.957, -0.604)
NormalizedLocationEntropy	0.451 (0.403, 0.546)	0.323 (0.253, 0.350)	0.128	0.640 (0.549, 0.732)
		Phone Usage		
SumDuration	400.204 (384.163, 416.432)	495.535 (480.862, 510.303)	-95.331	-0.535 [-0.758, -0.313]
CountEpisode	30.104 (25.562, 35.942)	40.645 (35.074, 45.521)	-10.541	-0.647 [-0.855, -0.439]
FirstUseAfter	28.745 [24.389, 33.867]	45.067 [41.483, 48.965]	-16.322	-0.578 [-0.739, -0.418]
MaxDuration	7.683 (6.422, 8.955)	18.073 (17.534, 19.131)	-10.390	-0.756 [-0.910, -0.603]
		Bluetooth		
UniqueDevices	3.701 (2.485, 4.933)	5.516 (4.433, 6.597)	-1.815	-0.238 (-0.302, -0.175)
CountScans	13.231 (11.630, 14.842)	19.096 (17.751, 20.448)	-5.865	-0.277 (-0.452, -0.102)
		Steps		
MaxSumSteps	5800.553 (5400.705, 6200.321)	6300.878 (5900.205, 6700.426)	-500.325	-0.518 (-0.604, -0.433)
AvgSumSteps	4800.335 (4500.634, 5100.074)	5300.745 (5000.832, 5600.642)	-500.410	-0.237 (-0.374, -0.128)
		Sleep		
AvgDurationAwake	60.320 (48.294, 72.041)	88.385 (78.037, 102.181)	-28.065	-0.451 (-0.647, -0.255)
AvgDurationaSleep	510.047 (405.378, 610.582)	407.731 (385.284, 433.539)	102.316	0.404 (0.308, 0.501)

Our analysis of location-based features found significant differences between socially lonely (SL) and emotionally lonely (EL) groups. The log-transformed location variance in the evening, which represents the variability in a participant's geographic position, was lower for the SL group (M = 2.301) compared to the EL group (M = 3.751), with a medium effect size of -0.715. This indicates that emotionally lonely individuals show more varied movement patterns in the evening. On the daily basis, the SL group visited fewer significant places, defined as distinct locations (M = 1.504), compared to the EL group (M = 2.167). The number of transitions between these significant locations was also lower for the SL group (M = 5.463) than the EL group (M = 7.374), with a large effect size of -0.780, indicating less movement between key locations for socially lonely individuals. Interestingly, the normalized location entropy, which measures the evenness of time distribution across significant locations, was higher for the SL group (M = 0.451) compared to the EL group (M = 0.323). This indicates that while socially lonely individuals visit fewer locations, they tend to distribute their time more evenly across these places.

Analysis of phone usage patterns also found significant differences between the two groups. The EL group showed higher overall phone engagement with a total daily unlock duration of 495.535 minutes compared to 400.204 minutes for the SL group (d = -0.535). The EL group took longer to first use their phone after waking, averaging 45.067 minutes versus 28.745 minutes for the SL group with an effect size of -0.578. Another difference was in the maximum duration of a single unlock episode, with the EL group spending up to 18.073 minutes compared to 7.683 minutes for the SL group. Other metrics, including the frequency of phone unlocks, also showed higher values for the EL group.

Analysis of Bluetooth related patterns also found significant differences between groups. It shows that the EL group encountered more unique devices daily (M = 5.516, 95) compared to the SL group (M = 3.701) which indicates increased social proximity or time spent in populated areas. Step count analysis showed the EL group had higher average daily steps (M = 5300.745) than the SL group (M = 4800.335), with a medium to large effect size (d = -0.754). Sleep patterns differed as well with the SL group sleeping longer on average (M = 510.047 minutes) compared to the EL group (M = 407.731 minutes). However, the EL group spent more time awake in bed (M = 88.385 minutes) than the SL group (M = 60.320 minutes) which might indicate sleep quality issues.

6.3.3 Predictive Power of Behavioural Features in Loneliness Categories

Table 6.3 presents the predictive performance of different ML classifiers for categorizing loneliness into 'socially lonely', 'emotionally lonely', 'both lonely', and 'not lonely'. We compared these classifiers against three baseline models: Majority Class (BL1: MC), Decision Tree (BL2: DT), and Random Weighted Classifier (BL3: RWC). These baseline models were chosen because they provide simple, interpretable benchmarks to evaluate the performance of more complex classifiers.

The XGBoost model achieved the highest overall accuracy of 78.48%. It outperformed in classifying the 'Both Lonely' category with an F1-score of 85.44% and showed strong performance in the 'Emotionally Lonely' category with an F1-score of 70.49%. The XGBoost model showed the best overall balance and highest metrics across all classes. After that, the Random Forest model achieved an accuracy of 75.58%, showing strong performance with an F1-score of 82.41% for the 'Both Lonely' class and 72.76% for the 'Not Lonely'class. The model provided a good balance of precision and sensitivity across all categories. Support Vector Machine

(SVM) performed well with 70.10% accuracy and made strong predictions in the 'Both Lonely' and 'Not Lonely' categories. K-Nearest Neighbors (KNN) shown balanced performance with a 65.53% accuracy.

While F1-scores provide a balanced performance measure but investigating precision and sensitivity separately can provide important observations regarding these models. For instance, the XGBoost model showed high precision (88.07%) in detecting students who are 'Both Lonely', which shows a low false positive rate for this category. However, its sensitivity for 'Socially Lonely' students was lower (58.59%), which shows some challenges in identifying all cases in this category. The Random Forest model showed a more balanced precision-sensitivity trade-off for the 'Not Lonely' category (precision: 75.94%, sensitivity: 70.86%) which shows consistent performance in both correctly identifying and capturing instances of this class. These nuances in precision and sensitivity across different loneliness categories highlight the varying challenges in accurately classifying each type of loneliness.

Table 6.3: Classification Results for Different Models on Loneliness Categories

Model	**Acc**	**Socially Lonely**			**Emotionally Lonely**			**Both Lonely**			**Not Lonely**		
		Prec	**Rec**	**F1**	**Prec**	**Sens**	**F1**	**Prec**	**Rec**	**F1**	**Prec**	**Sens**	**F1**
BL1: MC	42.54	0.00	0.00	0.00	0.00	0.00	0.00	42.54	100.00	59.59	0.00	0.00	0.00
BL2: DT	45.49	20.45	25.87	22.84	18.38	21.44	19.18	55.46	65.32	60.50	35.20	38.39	36.54
BL3: RWC	35.47	11.65	24.58	15.24	9.18	19.43	12.54	40.94	85.34	54.65	37.28	38.22	37.59
SVM	70.10	60.00	50.55	55.74	65.28	60.74	62.63	75.18	80.11	77.67	70.39	65.54	67.10
RF	75.58	65.43	55.74	60.54	70.15	65.24	67.08	80.84	85.72	82.41	75.94	70.86	72.76
KNN	65.53	55.11	45.69	50.64	60.48	55.00	57.06	70.15	75.75	72.38	65.57	60.92	62.14
XGBoost	78.48	68.37	58.59	63.62	68.63	73.27	70.49	88.07	83.13	85.44	73.58	78.52	75.27

6.3.4 Important Features for Loneliness Classification

Fig.6.1 and Fig.6.2 present the mean absolute SHAP values for different features across the four loneliness categories: Socially Lonely (SL), Emotionally Lonely (EL), Both Lonely (BL), and Not Lonely (NL). The x-axis represents the normalised mean absolute SHAP value for each feature, which shows the average magnitude of that feature's contribution to the model's output. SHAP values represent the impact of a feature on the model's prediction. The magnitude of a SHAP value shows the feature's importance for a particular prediction. In our figures, larger mean absolute SHAP values indicates a stronger average influence

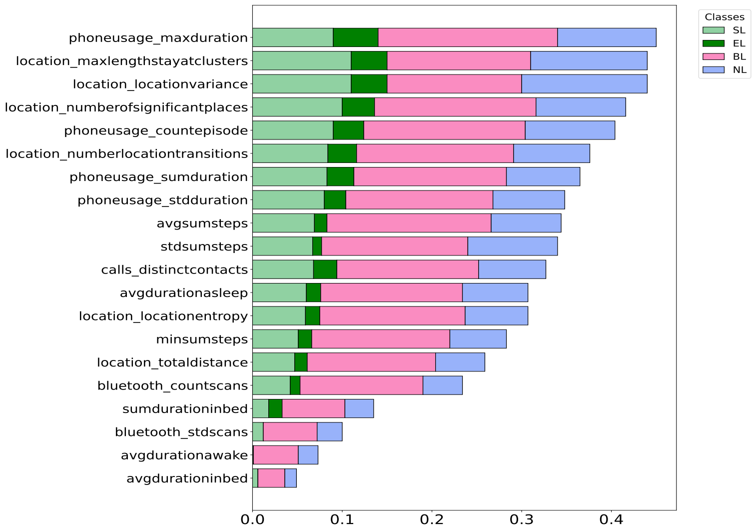


Figure 6.1: Mean absolute SHAP values for XGBoost model

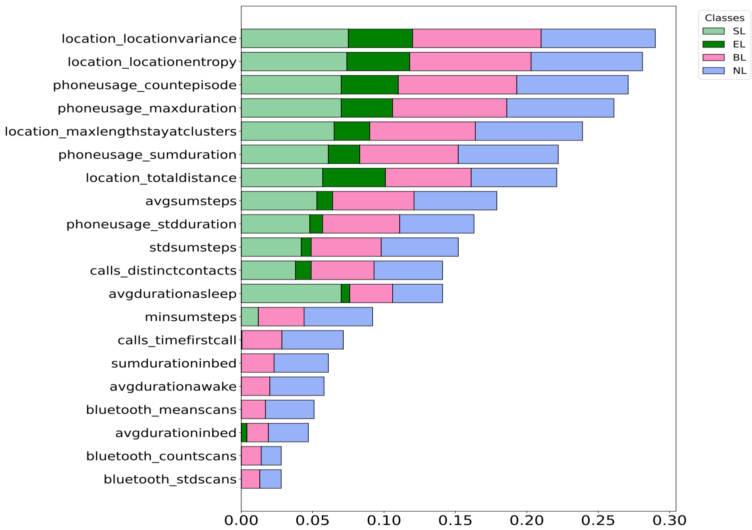


Figure 6.2: Mean absolute SHAP values for Random Forest model

of that feature on the model's decisions.

Both XGBoost and Random Forest models identified location-based and phone usage features as highly influential in distinguishing between loneliness categories. However, there were some differences in the relative importance of specific features between the two models. The most influential features in the XGBoost model were the maximum duration of phone usage and location maximum length of stay at clusters. These features showed a strong impact across all loneliness categories but the strongest effect was on the 'Both Lonely' category. Sleep-related features such as average duration awake also showed some importance specifically for the 'Not Lonely' category.

In our Random Forest model, the SHAP value analysis also found important findings related to features. Location-based features such as variance and entropy were highly influential and had the greatest impact on almost all categories. The maximum duration of the phone and location entropy were also influential. Interestingly, Bluetooth-related features appeared more important in the Random Forest model as compared to XGBoost.

6.4 Discussion

The main aim of this chapter was to explore how behavioural features from passive sensing data could distinguish between social and emotional loneliness and classify loneliness types using machine learning models. The analysis found statistically significant differences in the behavioural markers for social and emotional loneliness groups. The observed differences in behavioural features between the Socially Lonely (SL) and Emotionally Lonely (EL) groups provide evidence that passive sensing can capture distinct patterns associated with these two forms of loneliness.

Location-based behaviours showed significant differences between the two groups. The EL group showed higher location variance, especially in the evening, visited more significant places, and had more location transitions compared to the SL group. In contrast, the SL group's lower mobility and fewer significant places visited may indicate a lack of interest or opportunity to participate in social interactions, which is a sign of social loneliness [Rei21]. This behaviour might be linked to the Introversion-Extraversion Dimension of personality. Socially lonely individuals may lean towards introversion who prefer solitary activities and less social stimulation [Bur95]. There could also be some other factors for this behaviour, like a lack of social support, which might discourage them from exploring new environments or engaging in social activities. Interestingly, while the SL group visited fewer locations, they showed higher location entropy, which indicates more evenly distributed time across the places they visited. This might indicate a preference for familiar or comfortable environments, maybe somehow to manage feelings of social disconnection. This difference between the two groups shows the nuanced ways in which loneliness manifests. However, it is important to note that the current data cannot definitively establish a causal relationship. Existing studies provide some information about the reasons behind loneliness related behaviours [DJIHK18, LJF21, MRTALPCS21, HÖK^{+}22]. However a deeper investigation is still needed while using different methods like combining qualitative methods with behavioural features to gain a more nuanced understanding of the complex relationship between these two loneliness types and daily life behaviours.

The findings on phone usage patterns also show differences between the two types of loneliness. The EL group uses their phones more often and for longer periods, which might mean they prefer digital communication instead of face-to-face interactions. This could be explained by the idea of displacement hypothesis, where increased digital engagement takes the place of in-person social interactions and it could lead to increased feelings of isolation or loneliness [VP07]. The SL group uses their phones less, maybe because they have a different social situation or don't feel the need for digital contact.

The Bluetooth data shows more unique device encounters for the EL group, which supports the location data in suggesting that emotionally lonely individuals might seek out more populated areas. According to existing research, this behaviour aligns with the emotional loneliness construct, where individuals might feel lonely despite being in social settings [Wei75]. Significant differences in physical activity and sleep patterns between the two groups provide additional information about the behavioural impacts of loneliness. The SL group is less physically active which might cause or be a result of them withdrawing socially. They also sleep more, which could be due to many reasons like if they are feeling depressed or not active during the day, leading to a need for less sleep at night.

These findings have important implications for understanding the complex nature of loneliness and its impact on student behaviour and well-being. These findings show that passive sensing can differentiate between social and emotional loneliness and find the behavioural patterns associated with each. These findings can be used for targeted interventions, such as promoting social engagement for socially lonely students or addressing emotional needs and connectivity for emotionally lonely students. To maximize the effectiveness of such interventions, it is important to first get a deeper understanding of the motivations and experiences behind the behavioural differences we observed. More qualitative research with behaviour understanding could help us to understand why individuals engage in specific patterns and how they perceive the connection between their behaviours and feelings of loneliness.

The machine learning models trained for classification based on loneliness types showed their predictive capabilities of behavioural features extracted through passive sensing data. The best-performing model was XGBoost with an overall accuracy of 78.48%. This highlights the power of ensemble-based techniques for handling complex sensor datasets for loneliness prediction. Moreover the F1 score for XGBoost is also high for the 'Both Lonely' category as compared to other models. This shows the overall robustness of XGBoost for the classification of social and emotional loneliness.

The varying precision and sensitivity scores across different models and loneliness categories highlight important observations about the challenges of classifying loneliness types based on behavioural data. For instance, XGBoost's high precision (88.07%) for the 'Both Lonely' category indicates that when the model predicts this category, it's highly likely to be correct. This indicates that students experiencing both social and emotional loneliness may exhibit more distinct behavioural patterns that are easier for the model to identify. On the other hand, the lower sensitivity for the 'Socially Lonely' category across models indicates that behaviours associated with social loneliness might be more subtle or varied, making it challenging for models to identify all instances. This could imply that social loneliness manifests in more diverse ways in behaviours and might overlap with patterns seen in other loneliness categories, which makes it difficult for models to identify it correctly.

The feature importance analysis for XGBoost and Random Forest models provides important findings about the behavioural features that distinguish between different loneliness categories. The consistent importance of location-based and phone usage features across both models highlights their significance in understanding and categorizing loneliness. This indicates that mobility patterns play a significant role in differentiating loneliness types. Interestingly, our analysis found lower SHAP scores for features distinguishing between SL and EL groups compared to those differentiating the 'Both Lonely' and 'Not Lonely' categories. There could be multiple potential reasons for this observation. SL and EL might share similar behavioural patterns in some aspects which made it challenging for the models to distinguish between them based solely on passive sensing data. This means the behavioural manifestations of social versus emotional loneliness might be more nuanced and less pronounced in passive sensing data compared to the clear differences between being lonely (in any form) and not lonely. This could be because students experiencing both types of loneliness have shown more clear or consistent behaviours. At the same time, those who are not lonely have significant different patterns of movement, phone usage, and sleep.

6.4.1 Comparison with Chapter 5 Results

Chapter 5 and this chapter used the same GLOBEM dataset but in different contexts. Chapter 5 generic models were trained for binary classification with incremental learning while here we trained for multi-class classification using all the data at once. Comparing results from both chapters provides important observations about the complexities of loneliness detection using passive sensing data.

The XGBoost model consistently outperformed other algorithms in both chapters with an accuracy of 76.52% in Chapter 5 and 78.48% in this chapter. However, other metrics such as precision and sensitivity were lower in Chapter 5 for detecting lonely classes as compared to this chapter, 'both lonely' class. This could be due to inherent challenges with the incremental learning approach and benefits of batch learning in this chapter.

Both chapters showed strong performance in classifying general lonely vs not lonely which shows that passive sensing data can effectively capture overall loneliness states in behaviours. However, the multi-class model's lower performance in distinguishing between the SL and EL categories shows that these specific types of loneliness may have more subtle or overlapping behavioural manifestations which made it hard for the models to capture and learn effectively in the small sample. This also highlights the complexity of loneliness as a psychological construct and the challenges in its fine-grained classification based solely on behavioural data.

6.5 Conclusions and Limitations

In this chapter, we have explored the multifaceted nature of loneliness among students, specifically focusing on the distinction between social and emotional loneliness using passive sensing data. Our research has addressed three key questions that provide important insights into the behavioural manifestations of different types of loneliness. The main takeaways from this chapter are:

- Firstly, our analysis found that behavioural features extracted through passive sensing data can differentiate between socially and emotionally lonely students. The statistical test found statistically significant differences in various behavioural features between these two groups. Socially lonely individuals showed less variance in their locations as compared to emotionally lonely individuals. Additionally, socially lonely students had shorter overall phone usage duration, fewer Bluetooth scans, and fewer steps than emotionally lonely students.

- Our study has shown the considerable predictive power of behavioural features in classifying social and emotional loneliness. The XGBoost model achieved the highest overall accuracy of 78.48% and a high F1 score across all loneliness categories. This shows the potential of behavioural features extracted through passive sensing in identifying and differentiating between types of loneliness.

- We identified the most important behavioural features for predictive models in classifying loneliness and its types. Our analysis identified phone usage and location-based features as critical in distinguishing between loneliness categories in both XGBoost and Random Forest models. The findings highlighted that phone usage duration, location variance, and sleep-related features were particularly significant in differentiating socially, emotionally, and both socially and emotionally lonely individuals from those not experiencing loneliness.

Despite these significant findings, our study has some limitations that could be addressed in future research (see Chapter 8 for details):

- As we discussed in earlier chapters as well, the generalizability of our findings is limited by the size and diversity of the dataset. The participants, who are college students, form a homogeneous group with possibly similar daily routines and psychological challenges.

- The dataset we used does not capture all the characteristics associated with loneliness. Personal relationships [dJG89], age [SHGDF20], major life events [LEV20], and mental health history [HCH+22] all significantly influence an individual's experience of loneliness.

Chapter 7

The Relationship between Loneliness and Depression among College Students: Mining Data Derived from Passive Sensing

7.1 Overview

Throughout this thesis, we have explored the potential of passive sensing to find patterns of loneliness within the student population. We have presented how behavioural data can help us understand and detect the signs of loneliness. However, loneliness rarely exists in isolation; it is often deeply intertwined with depression. [PMHJ+21]. Loneliness is an experience in which a person perceives a lack of quality social relationships [PP81]. While Depression, or major depressive disorder (MDD), is a clinical mental health condition, characterized by constant feelings of sadness, despair, and lack of interest in engaging in activities [Pay08]. Within the broader context of student mental health, it is important to recognize the interconnection between loneliness and depression. These conditions often co-occur and can create a cycle of worsening emotional distress [LWGB15]. Loneliness can increase the risk of depression, and the two conditions can worsen each other [HG06, CHT10]. College students often go through new experiences and academic pressures and find themselves particularly vulnerable to these conditions [CBY+20]. There is a significant impact of loneliness and depression on the academic progress and well-being of students [LDB19]. While the connection between loneliness and depression is well-known in psychology literature, the understanding of the specific behavioural patterns that distinguish loneliness and depression is still limited [CHT10, QBR+13]. This understanding is crucial for developing targeted interventions that address the unique needs of students struggling with loneliness and depression. In this chapter, we aim to address this gap by using passive sensing data to identify unique behavioural markers associated with loneliness and depression in student population. This approach is novel, as no prior studies have used passive sensing data to study the behavioural connection between loneliness and depression.

In this chapter, we want to understand how behavioural patterns identified through passive sensing can reflect a person's experiences of loneliness, depression, or potentially even both. By filling this knowledge gap, we hope to improve the existing detection systems, which could lead to more personalized interventions that fit each person's specific needs. To investigate the relationship between loneliness, depression, and student behaviours, we used multiple methods, including regression analysis, mediation analysis, association rule mining (ARM), and machine learning. The key research questions we explored are:

1. What are the common and distinct behavioural patterns of students who are lonely or depressed, and how do these differ from those who are not?
2. How do loneliness and depression affect each other in the context of college students, and which specific behavioural markers serve as mediators in these relationships?
3. How effectively can behavioural markers, when combined with loneliness and depression scales as an input feature, identify lonely and depressed students?

7.2 Methods

7.2.1 Behavioural Patterns and their Relationship with Loneliness and Depression

In this study, we have used the StudentLife dataset. To answer our first research question related to finding common and distinct behavioural patterns between lonely and depressed students, we applied separate multivariate logistic regression models to independently model the probability of students experiencing loneliness and depression based on their behavioural features. Logistic regression was chosen for this analysis because it is particularly well-suited for addressing the research question of identifying behavioural patterns associated with loneliness and depression. It can model binary outcomes (lonely/not lonely, depressed/not depressed) to find a direct assessment of the probability of experiencing these conditions based on various behavioural features as predictors. It also provides interpretable coefficients to quantify the relationship between each behavioural feature and the likelihood of experiencing loneliness or depression. We used Mutual Information (MI) method to select most important features to include in this analysis.

The dependent variables in each model were two measures of loneliness and depression, specifically the UCLA loneliness scale [RPC80] and the PHQ-9 depression scale [KSW01]. The independent variables were behavioural features extracted from the StudentLife dataset. Our objective was to understand how changes in each independent variable (behavioural feature) affect the likelihood of being lonely and depressed. By modelling these relationships, we aimed to identify which behaviours are significantly associated with increased odds (positive association) or decreased odds (negative association) of experiencing loneliness or depression. The coefficients in each model represent the change in the log odds of the respective outcome (loneliness or depression) associated with a one-unit increase in the corresponding predictor variable (behavioural feature). These coefficients were transformed into odds ratios (OR) for easier interpretation. An OR greater than 1 indicates that a one-unit increase in the predictor is associated with an increase in the odds of the outcome, while an OR less than 1 indicates a decrease in the odds. For example, an OR of 1.5 for phone usage duration in the loneliness model would indicate that a one-unit increase in phone usage duration is associated with a 50% increase in the odds of being lonely. We considered results statistically significant if the p-value was less than 0.05. We used the bootstrap resampling method here to address the class imbalance and to obtain robust estimates of odds ratios (OR), p-values, and confidence intervals.

We also used Association Rule Mining (ARM) to complement the logistic regression analysis by looking into the combination of behavioural patterns associated with loneliness and depression separately. While logistic regression examines the independent effects of individual predictors, ARM finds how combinations of these features may collectively contribute to each outcome. This provides a more holistic understanding of the underlying behavioural mechanisms for both loneliness and depression. We used the Apriori algorithm [AIS93] which is a standard approach in ARM to discover frequent item sets and derive association rules from them. It can find complex combined patterns that might not be evident with logistic regression alone.

The Apriori algorithm is a two-step process for discovering patterns in transactional data such as the behavioural patterns of students in our study. It aims to uncover association rules that reveal relationships between different items (or behavioural categories) that frequently occur together.

Step 1: Identifying Frequent Itemsets

In the first step, Apriori identifies frequent itemsets. These are groups of items that appear together in a significant proportion of transactions. The support of an itemset is defined as the proportion of transactions that contain all items in the itemset. For example, if items A and B appear together in 30% of the transactions, then the support for the itemset A, B is 0.30.

Step 2: Generating Association Rules

Once the frequent itemsets are identified, Apriori proceeds to the second step which is generating association rules. An association rule is A $\rightarrow$ B, where A and B are itemsets. This rule implies that if itemset A is present in a transaction, then itemset B is also likely to be present. The strength of this association is measured by two metrics; confidence and lift. Confidence is the conditional probability of finding itemset B in a transaction given that itemset A is already present in that transaction. It's calculated as the ratio of the support of the combined itemset A, B to the support of itemset A alone. For instance, if the rule is A $\rightarrow$ B, and 90% of transactions containing A also contain B, the confidence of this rule would be 0.90. Lift measures how much more likely two items (or behaviours) are to occur together than if they were completely independent. A lift greater than 1 means these items tend to appear together more often than expected by pure chance which suggests a real connection or association between them.

The Apriori algorithm was run with a minimum support threshold of 0.1 (meaning that a pattern must appear in at least 10% of participants to be considered frequent), a minimum confidence threshold of 80% (meaning that the rule's antecedent must predict the consequent at least 80% of the time), and a minimum lift value of 1.5 (meaning that the association between the antecedent and consequent is 1.5 times stronger than expected if they were independent). To handle the varying scales of behavioural features, we categorized each into low, medium, and high groups based on their distribution, using the 33rd and 66th percentiles for categorization. This is a common practice in ARM, as it simplifies the analysis by reducing the complexity of the feature space and it provides more interpretable association rules. Additionally, this categorization aligns with our research question, which aims to find meaningful patterns of behaviour rather than precise numerical values.

7.2.2 Mediation Analysis: Exploring the Relationship between Loneliness and Depression

We used mediation analysis to investigate how loneliness and depression influence each other through different behavioural features. Mediation analysis is a statistical method that helps to identify how intermediate variables or behaviours mediate the relationship between 2 variables [MCP12]. We used this approach to address the second research question about exploring the relationship between loneliness, depression, and behavioural features and identifying specific behaviours that may serve as pathways through which these two conditions influence each other.

Mediation analysis involves examining how an independent variable (IV) affects a dependent variable (DV) through a mediator variable (M) which can be seen in Figure 7.1. IV is the initial variable that we hypothesize associated with the DV. M is the intermediate variable through which the IV influences the DV. While DV is the outcome variable that we are interested in. Our focus is on the indirect effect, which represents the pathway through the mediator.

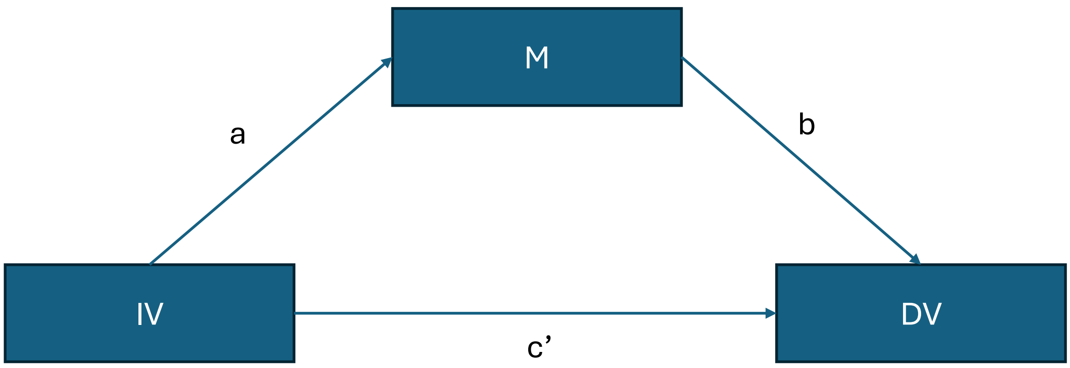


Figure 7.1: Mediation Analysis For Relationship Between an Independent Variable (IV), Mediator (M), and Dependent Variable (DV). The direct effect (c’) represents the relationship between the IV and DV, while the indirect effect (a*b) represents the effect of the IV on the DV through the mediator.

The analysis involves two steps to compute the indirect effect. First, regress the mediator (M) on the independent variable (X) to obtain the coefficient α.

$$M = \alpha_0 + \alpha_1 X$$

Here, α_0 is the intercept term, representing the expected value of M when X is zero. α_1 represents the slope, indicating how much M changes for a one-unit increase in X. Together, these coefficients describe the linear relationship between X and M.

Next, regress the dependent variable (Y) on both the independent variable (X) and the mediator (M) to obtain the coefficients β_1 (for the mediator) and β_2 (for the independent variable).

$$Y = \beta_0 + \beta_1 M + \beta_2 X$$

The indirect effect is calculated as α_1 x β_1, representing the effect of X on Y through the mediator.

To investigate the effect of loneliness on depression through behavioural features, We conducted mediation analyses using loneliness as the independent variable (IV) and depression as the dependent variable (DV). Similarly, while studying the effect of depression on loneliness, depression served as the IV and loneliness as the DV. Each behavioural feature was investigated as a potential mediator variable. We used the bootstrapping method with 10k samples to test the significance due to its robustness with small sample sizes. Based on EDA analysis, only those features were included in this analysis that show linear relationship with the IV and DV as it is an assumption for mediation analysis.

7.2.3 Machine Learning Prediction of Loneliness and Depression

To assess the efficacy of behavioural features in predicting loneliness and depression, we developed two separate sets of binary classification models: one for loneliness and another for depression. For each classification task, we used multiple algorithms, including logistic regression, k-nearest neighbours, support vector machine, random forest, and XGBoost. For the loneliness classification models, we used behavioural features in combination with the PHQ-9 depression scale scores as predictors. Similarly, for the depression classification models, we used the same behavioural features along with the UCLA loneliness scale scores as predictors.

Our model evaluation used the pipeline with nested cross-validation method. In the outer loop, we implemented leave-one-subject-out cross-validation to evaluate generalization. For each iteration of the outer loop, we have an inner loop of stratified three-fold cross-validation for hyperparameter tuning using GridSearchCV. Preprocessing steps, including missing values imputation, feature scaling, feature selection and handling class imbalance using SMOTE were performed within each inner fold before model training and evaluation. This approach minimized the risk of overfitting by tuning hyperparameters on multiple stratified subsets of the training data.

For baseline comparisons in loneliness prediction, we considered the majority class, random weighted classifier, and a decision tree model that used only the PHQ-9 Depression Scale score as an input feature. Similarly, for depression prediction baselines, we used the majority class, random weighted classifier, and a decision tree model with only the UCLA Loneliness Scale score as the input feature. These baselines helped us to assess the added value of adding behavioural features in our more complex models. Performance metrics such as accuracy, precision, sensitivity, F1-score, and area under the receiver operating characteristic curve (AUC) were calculated for each model.

7.3 Results

7.3.1 Identified Behavioural Patterns of Lonely and Depressed Students

The regression analysis findings (Table 7.1) show significant associations between behavioural features and loneliness and depression. The following sections present the key findings for each behavioural feature.

Phone Usage The phone usage features showed several significant associations with loneliness and depression.

- Loneliness: We found a strong positive association between overall phone usage duration and loneliness. Students who spent more time on their phones throughout the day had higher odds of experiencing loneliness (odds ratio of 1.50, p=0.003). This connection remained significant even when

Table 7.1: Results of Regression Analysis for Loneliness and Depression: Odds Ratios, 95% Confidence Intervals, and p-values for Behavioural Features

behavioural Feature	Loneliness			Depression		
	Odds Ratio	**95% CI**	**p-value**	**Odds Ratio**	**95% CI**	**p-value**
Total Number of screens unlocks	1.20	[1.05,1.37]	0.031	0.75	[0.62,0.91]	0.044
Number of screens unlocks (evening)	1.75	[1.32,2.31]	0.001	0.80	[0.65,0.95]	0.021
Phone usage duration	1.50	[1.20,1.88]	0.003	0.70	[0.52,0.94]	0.018
Phone usage duration (evening)	1.45	[1.15,1.85]	0.002	1.10	[0.95,1.26]	0.310
Number of location transitions	1.73	[1.30,2.20]	0.016	0.65	[0.50,0.95]	0.034
Number of unique places visited	1.13	[1.03,1.23]	0.021	0.85	[0.76,0.95]	0.023
Activity duration	0.72	[0.54,0.91]	0.014	0.69	[0.55,0.87]	0.022
Sleep duration	0.46	[0.30,0.70]	0.002	0.75	[0.62,0.90]	0.003
Number of conversations	1.50	[0.95,2.38]	0.101	0.75	[0.63,0.89]	0.002
Number of incoming calls	0.82	[0.71,0.94]	0.023	1.10	[0.90,1.35]	0.351
Number of incoming calls (evening)	0.30	[0.20,0.45]	0.002	1.10	[0.95,1.26]	0.303

focusing specifically on evening phone usage (odds ratio of 1.45, p=0.002). Similarly, features related to the number of screen unlocks also show a positive association with loneliness.

- Depression: The trend for depression was different. Higher phone usage was associated with lower odds of depression (odds ratio of 0.70, p=0.018). However, the relationship between evening phone usage and depression was not statistically significant. Similarly, the screen unlock features showed a negative association with depression.

Locations The location features include the number of significant locations visited (calculated using clustering algorithms to identify clusters, where each cluster represents a significant place) and the number of location transitions (the number of movements between any two clusters in a time segment) [VLA+20].

- Loneliness: Students who have a higher number of location transitions overall throughout the day were more likely to experience loneliness (odds ratio of 1.73, p=0.016). However, this association weakened somewhat when considering the number of unique places visited (odds ratio of 1.13, p=0.021).

- Depression: Interestingly, we found the opposite trend for depression. Students with a high number of location transitions, particularly unique ones, had lower odds of experiencing depression (odds ratio of 0.65, p=0.034 for overall transitions and odds ratio of 0.85, p=0.023 for unique places).

Physical Activity Duration Higher physical activity duration was associated with a reduced likelihood of both loneliness and depression. The odds ratios were almost similar for both conditions (0.72, p=0.014) for loneliness, and for depression, 0.69, p=0.022.

Sleep Just like physical activity, sleep duration was negatively associated with both loneliness and depression. The odds ratio for loneliness is 0.46 with p=0.002 and for depression, the odds ratio is 0.75 with p=0.003.

- Loneliness: Longer sleep duration was associated with a 54% decrease in the likelihood of feeling lonely (odds ratio of 0.46, p=0.002).
- Depression: Similarly, students who slept longer had lower odds of experiencing depression (odds ratio of 0.75, p=0.003).

Conversation While number of conversations was associated with a lower likelihood of depression (odds ratio of 0.75, p=0.002), this relationship was not statistically significant for loneliness.

Calls Our analysis found that the number of incoming calls may play a role in reducing feelings of loneliness. Students with more incoming calls throughout the day had lower odds of experiencing loneliness, with an odds ratio of 0.82. Similarly, the number of incoming calls in the evening showed an odds ratio of 0.30 for loneliness. However, the relationships between the number of calls and depression, both during the day and in the evening were not found to be statistically significant.

Association Rule Mining (ARM) Findings

As presented in Table 7.2 our ARM analysis found a combination of behavioural patterns associated with loneliness. The combination of behavioural patterns with low location transitions, low activity duration, high overall screen unlocks, low sleep, and medium incoming calls was strongly associated with loneliness,

Table 7.2: Results of Association Rule Mining for Loneliness: Support, Confidence, Lift, and Behavioural Patterns

Rule	behavioural Patterns	Support	Confidence	Lift
1	Low location transitions, low activity duration, high total screen unlocks, low sleep, medium incoming calls, low day Bluetooth scans	0.14	0.84	1.63
2	High phone usage, medium total screen unlocks, high evening screen unlocks, low sleep, low conversations, low day Bluetooth scans	0.13	0.88	1.71
3	High evening phone usage, high evening screen unlocks, low conversations, low activity duration, low sleep, low evening incoming calls	0.12	0.85	1.65
4	High total screen unlocks, high unique places visited, low sleep, low incoming calls, low activity duration, medium conversations	0.11	0.82	1.59
5	Low location transitions, low unique places visited, high evening phone usage, low sleep, low conversations, high evening screen unlocks	0.10	0.81	1.57

being observed in 14% of students with 84% confidence. Similarly, students with high evening phone usage, high evening screen unlocks, low conversations, low activity duration, low sleep, and low evening incoming calls had a strong positive association with loneliness. This pattern was present in 12% of the students, with a high confidence of 85% and a lift of 1.65 which shows that this combination of behaviours increases the likelihood of experiencing loneliness. Additionally, high total screen unlocks, high unique places visited, low sleep, low incoming calls, low activity duration, and medium conversations were found in 11% of the students, with an 82% confidence and a lift of 1.59 a similar. Another important pattern included high phone usage, medium total screen unlocks, high evening screen unlocks, low sleep, low conversations, and low day Bluetooth scans, present in 13% of the students with a confidence of 88% and a lift of 1.71.

Similarly, Table 7.3 presents patterns found from ARM for depression. The most

Table 7.3: Results of Association Rule Mining for Depression: Support, Confidence, Lift, and Behavioural Patterns

Rule	behavioural Patterns	Support	Confidence	Lift
1	Low location transitions, low unique places visited, medium phone usage, high evening screen unlocks, low activity duration, medium incoming calls	0.14	0.81	1.58
2	Low sleep, high total screen unlocks, low incoming calls, low conversations, low activity duration, high evening screen unlocks	0.12	0.86	1.67
3	High screen unlocks, medium phone usage, low location transitions, low unique places visited, low activity duration, low day Bluetooth scans	0.11	0.83	1.62
4	Low sleep, low conversations, high total screen unlocks, low incoming calls, low activity duration, high evening screen unlocks	0.10	0.80	1.55

prevalent pattern, occurring in 14% of students with 81% confidence included low mobility, moderate phone usage, low physical activity, and moderate social interaction. Another strong pattern, present in 12% of students with 86% confidence and the highest lift of 1.67, combined poor sleep, high screen unlocks, and low physical activity.

7.3.2 How Loneliness and Depression Affect Each Other: Findings from Mediation Analysis

In this section, we present the results of the mediation analysis to explore the effects of various behavioural features on the relationship between loneliness and depression among students. The analysis aims to identify which behaviours mediate the impact of loneliness on depression and vice versa.

We found a significant negative indirect effect of loneliness on depression through activity duration (Estimate = -0.72, p = 0.014). This shows that increased loneliness is associated with decreased activity duration, which in turn is linked to higher levels of depression. Similarly, sleep duration also showed a negative effect (Estimate = -0.50, p = 0.002). The indirect effect of the number of conversations on the relationship between loneliness and depression was significant (Estimate = -1.07, p = 0.044). This means that increased loneliness is associated with a decrease in the number of conversations, which in turn is linked to higher levels of depression. The number of location transition features did not show a significant estimate for loneliness on depression.

Table 7.4: Mediation Analysis Results for Loneliness and Depression

	Loneliness → Depression			Depression → Loneliness		
Variable	**Estimate**	**p-value**	**95 CI**	**Estimate**	**p-value**	**95 CI**
activity duration	-0.72	0.014	(-0.91, -0.54)	-0.71	0.022	(-0.87, -0.55)
sleep duration	-0.50	0.002	(-0.70, -0.30)	-0.76	0.003	(-0.90, -0.62)
phone usage duration during evening	-1.10	0.310	(-2.50, 0.30)	1.50	0.002	(1.15, 1.85)
number of location transitions	1.73	0.116	(-0.40, 3.86)	0.72	0.034	(0.50, 0.95)
number of conversations	-1.07	0.044	(-1.14, -1.01)	-1.11	0.250	(-2.10, 0.12)
number of conversations during day	1.40	0.121	(-0.36, 3.16)	-0.77	0.004	(-0.92, -0.62)

For the effect of depression on loneliness through mediating variables, we found a positive indirect effect of through evening phone usage duration (Estimate = 1.50, p = 0.002). This shows that as evening phone usage increases, the effect of depression on loneliness also increases. The number of location transitions also showed a positive indirect effect (Estimate = 0.72, p = 0.034). This means that an increased number of location transitions significantly mediates the effect of depression on loneliness. Activity duration and sleep duration showed negative effects, which means higher levels of depression are associated with decreased activity and sleep duration, which in turn are linked to increased feelings of loneliness.

7.3.3 Machine Learning Predictive Power for Loneliness and Depression

We explored the potential of various machine learning models to predict loneliness and depression. The performance of these models is evaluated based on several metrics, including accuracy, area under the curve (AUC) and F1 scores for both classes (1 and 0).

Table 7.5 shows the performance of different classification models for predicting loneliness. All models outperformed the baseline models. The XGBoost model achieved the highest accuracy of 82.43%, with an AUC of 83.31%. For the lonely class (class 1), XGBoost showed a precision of 70.97%, sensitivity of 60.49%, and F1 score of 65.36%. For the non-lonely class (class 0), it showed a precision of 84.17%, sensitivity of 91.28%, and F1 score of 87.18%. In comparison, the Support Vector Machine (SVM) model came in second, with an accuracy of 76.90%, an AUC of 75.89%. The Random Forest model ranked third, with an accuracy of 70.82%, an AUC of 68.08%. The K-nearest Neighbors, Logistic Regression showed relatively lower performance across all metrics.

Table 7.6 presents the performance of the classification models for depression prediction. Similar to the loneliness prediction results, the XGBoost model outperformed all other models, achieving an accuracy of 79.43%, an AUC of 80.21%. For the depressed class (class 1), XGBoost showed a precision of 65.87%, sensitivity of 56.39%, and F1 score of 60.86%. For the non-depressed class (class 0), it demonstrated higher performance with a precision of 83.99%, sensitivity of 89.25%, and F1 score of 86.52%. The Support Vector Machine (SVM) model ranked second, with an accuracy of 74.60%, an AUC of 73.49%. Random Forest showed an accuracy of 67.72% with AUC of 65.98%.

Table 7.5: Performance of Classification Models for Loneliness Prediction. The model performances are compared with three baseline classifiers (MC Majority Class, RWC Random Weighted Classifier, DT Decision Tree). All values are reported as percentages.

Models	Accuracy	AUC	Precision1	Sensitivity1	F11	Precision0	Sensitivity0	F10
Baseline1: MC	53.05	50.00	0.00	0.00	0.00	63.78	100.00	86.39
Baseline2: DT	51.53	47.48	26.29	20.38	23.26	65.58	82.18	71.39
Baseline3: RWC	58.72	52.64	32.48	24.21	28.64	71.39	76.39	74.82
Logistic regression	66.84	70.26	49.73	58.82	54.92	80.18	73.19	77.33
Support vector machine	76.90	75.89	58.25	54.77	56.36	84.54	86.48	85.48
K-nearest neighbours	68.25	69.35	42.93	65.76	51.95	85.12	69.13	76.30
Random forest	70.82	68.08	44.06	43.84	43.95	80.21	80.35	80.28
XGBoost	82.43	83.31	70.97	60.49	65.36	84.17	91.28	87.18

Table 7.6: Performance of Classification Models for Depression Prediction. The model performances are compared with three baseline classifiers (MC Majority Class, RWC Random Weighted Classifier, DT Decision Tree). All values are reported as percentages.

Models	Accuracy	AUC	Precision1	Sensitivity1	F11	Precision0	Sensitivity0	F10
Baseline1: MC	69.07	50.00	0.00	0.00	0.00	69.07	100.00	81.61
Baseline2: DT	56.26	45.96	18.43	19.43	18.93	69.50	68.50	69.00
Baseline3: RWC	58.68	48.13	23.07	24.14	23.75	70.15	71.12	70.61
Logistic regression	67.81	66.21	45.61	52.11	48.68	79.26	74.86	76.93
Support vector machine	74.60	73.49	53.15	47.67	50.26	81.54	83.38	82.43
K-nearest neighbours	65.15	66.25	39.83	58.66	47.85	82.02	67.03	73.90
Random forest	67.72	65.98	41.96	40.74	41.35	77.11	78.25	77.68
XGBoost	79.43	80.21	65.87	56.39	60.86	83.99	89.25	86.52

7.4 Discussion

Our study found important aspects of the complex relationships between behavioural patterns, loneliness and depression. While these conditions often co-occur, our findings show that they may stem from different underlying mechanisms and are associated with distinct behavioural patterns.

Firstly, we found several behaviours that appear to have the same relationship with both loneliness and depression. For example, we saw that more physical activity throughout the day was linked to lower feelings of loneliness and depression. This means that staying active might help protect against both conditions. Similarly, longer sleep durations were also associated with lower loneliness and depression. These results are aligned with previous research and highlight the role of sleep and physical activity for better mental well-being [CHC+02, KKH+11, JTB+20].

However, this chapter also found interesting differences in how certain behaviours relate to loneliness versus depression. For example, phone usage patterns showed a complex and apparently contradictory relationship between these two conditions. We found that increased phone usage was associated with higher levels of loneliness but lower levels of depression. This difference could come from how individuals use their phones in response to their emotional states. In Chapter 2, we discussed that studies have found that students with loneliness tend to spend more time on their phones, and loneliness was shown to be positively linked to increased phone usage [PTAK16, DVC+19, LSWL16]. Regarding the relation of phone usage with depression, some existing studies support our findings. For instance, a study discovered that greater smartphone use was related to lower depression severity [ETW+18]. They propose that phone use might provide mood enhancement or social connection that could reduce depressive symptoms. Similarly, another study found that while screen time minutes were not related to depression and anxiety severity, the frequency of phone screen unlocking (indirect proxy of phone usage) was negatively correlated with these conditions [RLHE18]. However, some other studies have found a positive association between phone usage and depression [AAT+18, YFLL20]. These conflicting findings on the link between phone usage and depression are likely due to differences in study populations, recall bias, usage patterns, and the complex nature of this relationship as explained in [ASM+24].

Increased phone usage might be associated with fewer face-to-face interactions, which could be linked to higher feelings of loneliness. As individuals spend more time on their phones, they may miss out on meaningful social interactions that could alleviate loneliness. This aligns with previous studies showing a positive link between loneliness and increased phone usage among students [PTAK16, LSWL16]. On the other hand, for depressed people phone usage might serve as a distraction or a coping mechanism. Engaging with online content or communicating with others, even if superficially, can provide temporary relief from depressive symptoms. This explanation is supported by [CKC+15], who found a negative association between depression scores and smartphone addiction scores among college students. Another study further explained this complex relationship and indicated that increased phone use might serve as a temporary stress relief mechanism that could lead to lower depressive symptoms [SNČS21].

Another observation shows that visiting a greater number of places throughout the day was associated with increased loneliness, which aligns with findings of an existing study in individuals with schizophrenia where fewer significant locations visited correlated with higher loneliness, potentially due to differing types of locations visited and resulting social interactions [FMG+21]. People who visit more places might be doing so in an attempt to cope with their loneliness through superficial social interactions, which may not provide meaningful connections and could lead to reinforcing their feelings. This interpretation is supported by a study finding that while interactions with weak ties can increase happiness, they don't necessarily reduce loneliness [SD14].

However, there was a negative association between the number of places visited and depression, which aligns with research indicating that increased physical activity and engagement are linked with reduced depressive symptoms [MTOA+21]. Another recent study that used AI models to predict depression risk based on mobility patterns found that higher mobility, location entropy, and GPS samples in transition was associated with a lower risk of depression [ASM+24]. However, it is important to consider that the relationship between location patterns and mental health outcomes may vary across different demographic and socioeconomic groups. The same study also found that while the negative association between location entropy and depression risk held true for majority and higher socioeconomic subgroups, the opposite relationship was observed in lower socioeconomic subgroups. For instance, higher location entropy decreased depression risk for those from middle-income families but increased risk for those from low-income backgrounds [ASM+24]. This explains the observed contradictory findings, and it also highlights the importance of considering other demographic variables as well to get the true picture.

Our observation from Chapter 6 that emotionally lonely individuals show higher mobility and more varied location patterns aligns with the finding in this chapter that an increased number of places visited is associated with higher loneliness but lower depression. This supports the idea that emotionally lonely individuals might engage in more social-seeking behaviours potentially as a coping mechanism, which might temporarily reduce depressive symptoms but not necessarily reduce feelings of loneliness. Similarly, the contrasting relationships of phone usage with loneliness and depression observed in this chapter provide context for the higher phone engagement seen in emotionally lonely individuals in Chapter 6. This suggests that while increased phone usage might be an attempt to cope with emotional loneliness, its effectiveness in reducing loneliness versus depressive symptoms may differ. These interconnected findings highlight the importance of distinguishing between different types of loneliness and their unique relationships with behavioural patterns and mental health outcomes.

Our ARM analysis provides important findings that complement our regression results. While logistic regression identified individual risk factors, ARM found how these factors interact and co-occur, which gives a more nuanced understanding of the behavioural contexts associated with loneliness and depression. For both loneliness and depression, we identified several recurring behavioural markers. Low sleep duration, high screen unlocks (particularly in the evening), and low physical activity consistently appeared in patterns associated with both conditions. This aligns with existing literature on the importance of sleep and physical activity in mental well-being [SNČS21, THH11, JTB+20]. However, we also observed distinct patterns for each condition. Loneliness was prominently associated with high evening phone usage, while depression patterns more consistently included low mobility (low location transitions and unique places visited). Similarly, mediation analysis found complex and sometimes bidirectional relationships between loneliness, depression, and various behavioural features. Some findings, such as the negative indirect effects of activity and sleep duration, reinforced our earlier results. While some others provided unique aspects of the specific pathways through which these factors interrelate. For instance the significant positive indirect effect of evening phone usage from depression to loneliness, but not in the reverse direction. Similarly, location transitions showed a significant effect only in the depression-to-loneliness pathway. This analysis highlights the importance of considering directional influences.

Our machine learning analysis further proved the potential of behavioural features in detecting loneliness and depression among students. The XGBoost model consistently outperformed other models for both loneliness and depression prediction. Interestingly, the models showed slightly better performance in predicting loneliness compared to depression, which may reflect the more direct relationship between observable behaviours and feelings of social isolation, but this needs to be confirmed with bigger datasets.

Another important observation is about the performance difference between majority and minority classes in both loneliness and depression prediction. This persistent imbalance highlights that the underlying patterns distinguishing lonely or depressed students from not lonely/depressed are inherently complex. The higher performance metrics for the majority classes (non-lonely and non-depressed) even after applying SMOTE indicate that it's generally easier to identify students not experiencing these conditions. While the models show good overall performance but this lower precision and sensitivity for the minority classes highlights the need for carefully using such models in clinical applications.

In this chapter, we applied Leave-One-Subject-Out (LOSO) cross-validation for the loneliness classification using the StudentLife dataset, whereas in Chapter 4, we used 10-fold cross-validation. The key difference between these two approaches lies in how the data is split for training and testing. 10-fold cross-validation, used in Chapter 4, splits the entire dataset into 10 equal parts which ensures a relatively balanced distribution of positive (lonely) and negative (not lonely) samples across the folds. This results in more evenly distributed test sets and contributes to more stable performance metrics. However, a key limitation of this 10-fold CV approach is that it might allow the model to see data from each student multiple times across different folds. This could lead to a subtle form of data leakage, as the model might indirectly learn individual patterns across folds, which may artificially improve performance metrics.

On the other hand, LOSO cross-validation applied in this chapter isolates one student's data entirely for testing in each iteration, while the remaining students' data is used for training. This approach ensures that the model does not see any part of the test student's data during training to get a strict evaluation of the model's generalizability to new subjects. However, this can lead to imbalance in the test set, especially when individual students show behaviours that are not well-represented in the training data. This imbalance can cause the model to struggle to predict loneliness for students whose behavioural patterns are under-represented in the training set. This test set imbalance and elimination of data leakage in this chapter result in lower performance metrics compared to Chapter 4 results. Future work should consider applying LOSO in Chapter 4 to ensure a fair comparison across approaches.

Collectively all these findings directly address our core research questions which aimed to understand the interconnected relationship between behaviours extracted through passive sensing with loneliness and depression. By investigating these associations, we have gained a holistic picture of the specific factors that contribute to these two conditions among students. Our primary aim was to explore the role of behavioural patterns in the context of loneliness and depression. As discussed in Chapter 1 and 2 that existing research has primarily examined these features in relation to general mental health conditions, we have shown that these same features interact in distinct ways when examining loneliness versus depression. The findings of this chapter support and expand upon the previous literature reviewed in Chapter 2, which emphasized the role of behavioural markers in mental health detection. While our study confirms the importance of features like phone usage, sleep duration, and physical activity, our findings highlight the need for a more nuanced interpretation of these features, particularly when distinguishing between loneliness and depression. The results presented in this chapter show that while passive sensing features are powerful tools for detecting mental health issues, they must be interpreted in context to accurately differentiate between specific conditions like loneliness and depression. This highlights the need of further research into how contextual and emotional states modulate these behavioural patterns differently in specific mental health conditions.

While our analyses provides important aspects about the relationship between loneliness and depression and the potential mediating role of behaviours but it does not establish causality between loneliness and depression. The observed relationships should be interpreted as associations that need further investigation through longitudinal or experimental studies to confirm the directionality and causal nature of these effects.

7.5 Conclusions and Limitations

This chapter investigated the relationship between loneliness and depression using passive sensing data. Below are the key findings from this chapter that address the key research questions presented at the beginning of this chapter:

- We have found distinct behavioural patterns associated with loneliness and depression. Higher phone usage was associated with increased loneliness but decreased depression. Location-based features showed opposite trends for loneliness and depression, with more location transitions associated with higher loneliness but lower depression. Physical activity and sleep duration were positively associated with reduced likelihood of both conditions. These findings highlight the nuanced differences in behavioural manifestations of loneliness and depression that emphasize the importance of considering these conditions separately in intervention strategies.

- The mediation analysis found significant indirect effects between loneliness and depression through different behavioural features. Phone usage duration, number of unique places visited, physical activity duration, and sleep duration were identified as significant mediators in both directions. These findings show a nuanced relation between loneliness and depression, mediated by different behaviours. For instance, while increased phone usage might increase the effect of depression on loneliness, it appears to mitigate the impact of loneliness on depression.

- We have shown the high predictive power of passive sensing data in identifying students experiencing loneliness or depression. The XGBoost model achieved the highest accuracy for both loneliness (82.43%) and depression (79.43%) predictions. This shows that behavioural features derived from passive sensing data, combined with standardized scales can effectively identify lonely or depressed students. The strong performance of these models highlights the potential of passive sensing in early detection and intervention strategies for student mental health.

We will be discussing detailed limitations along with future work in the next chapter, but below are some key limitations that we want to acknowledge:

- As mentioned in earlier chapters as well, these findings are limited to a single student population. Therefore, these findings should be interpreted within this context and not generalized to other populations or broader age groups.

- The interpretation of behavioural features (e.g., phone usage) may vary among individuals. For instance, high phone usage might indicate social engagement for some but signify isolation for others as we have explained in the discussion section.

- While significant associations were found between loneliness, depression and behavioural features, our study did not infer causality. Longitudinal studies or experimental designs would be necessary to establish causal relationships and better understand the directional influences between these conditions through behavioural patterns.

Chapter 8

Conclusion and Future Work

This chapter presents the key findings of this thesis. It also presents some limitations of this work and potential areas for future research that could extend and refine the work presented in this thesis.

8.1 Conclusion

Loneliness is becoming a significant worldwide issue with physical, emotional, social, and economic consequences. Multiple aspects, such as age, gender, socioeconomic level, and significant life events, could impact this complex condition. Early detection of loneliness is important to avoid its harmful long-term physical and mental health effects. Various factors present challenges to the early detection of loneliness, including its subjective nature, the stigma associated with loneliness, less self-awareness, etc. The younger population is more vulnerable to loneliness due to certain factors associated with this specific population. The challenges of early detection and potential mental health stigma further complicate early detection in this population group. Fortunately, the younger population is more comfortable and literate in using technology. This makes using smartphones and wearable devices for early loneliness detection in this population specifically.

Chapter 2 of this thesis presents a comprehensive scoping review to analyse the existing research on loneliness and social isolation detection using passive sensing. This literature review is the first to look at this specific area of loneliness detection with passive sensing. The main aim of the review was to synthesize the existing body of research, identify trends and limitations, and propose potential directions for future research. This review provided an important overview of the strengths and limitations of passive sensing techniques used in loneliness and social isolation detection. The analysis found the increasing use of smartphones and wearable sensors for loneliness detection across various studies and demographics. The review also highlighted several gaps in existing studies on loneliness detection, specifically privacy aspects of such systems, reliance on generic machine learning models, lack of dynamic analysis, energy consumption-related issues, and clinical validation issues of these systems.

Chapter 3 presents the methodology of our research. It provided a comprehensive introduction to the datasets used, including specific sensor details. The chapter elaborated on the extraction process of behavioural features, followed by a thorough explanation of data-cleaning steps. Additionally, it described the machine learning models used.

Building upon the gaps identified in Chapter 2, Chapter 4 presented the personalisation-based approach for loneliness detection using student grouping based on behavioural similarities. Since loneliness is a very subjective experience which can manifest in people very differently, existing methods in literature often overlook the crucial individual differences among people. This makes it difficult for generic models to accurately predict its presence and severity. A more effective approach could be identifying subgroups with similar behavioural characteristics, which would lead to a more personalized and accurate detection of loneliness and the behavioural patterns within each subgroup. Based on this idea, we used clustering techniques on the StudentLife dataset to identify distinct behavioural subgroups to develop personalized loneliness detection models. We also monitored the evolution of these groups throughout the study period and applied binary classification algorithms to each group to detect loneliness. Our results show that this categorization of students based on behavioural patterns increases the performance of loneliness detection models over a generic approach. This work presents a novel perspective on loneliness detection in two ways. First, it addresses the limitations of generic models by recognizing the diversity of how loneliness manifests in individual behaviours. By identifying subgroups based on similar behavioural patterns, the models can be customized for better performance. Secondly, this study presents the idea of tracking behavioural changes over time for loneliness detection. This approach gives useful information into how loneliness may emerge or intensify in individuals which could potentially be early warning signs.

Building upon the personalized approach introduced in Chapter 4, we recognized the need for a dynamic approach that could adapt to the changing behavioural patterns of students and can detect loneliness on an early basis. In Chapter 5, we presented a novel dynamic approach that identifies diverse behavioural subgroups, tracks their dynamic evolution, analyzes their relationship with loneliness and detects loneliness in a dynamic way. For this, we used the larger GLOBEM dataset. This approach updates and refines the behavioural groups based on new data on a weekly basis and in parallel also updates the classification models for loneliness detection. This dynamic approach ensures that our models keep up with the changing nature of student life, which may lead to more accurate and timely loneliness detection. To address the ever-changing nature of student behaviours, we used incremental clustering to update behavioural subgroups as new data arrives. We then developed both a broad generic classification model and specialized models for each subgroup. This approach recognizes that loneliness manifests differently across individuals. Our results show that student behaviour changes over time. We saw how groups of students with similar behaviours shifted, and how this was linked to feelings of loneliness. This highlights the need for detection systems that include the ability to evolve and adjust to reflect the dynamic nature of students' lives.

This thesis extends beyond detecting loneliness as a single concept. Researchers have identified two facets of loneliness that can manifest in individuals: social and emotional loneliness. In Chapter 6, we investigated the manifestation of social and emotional loneliness in behaviours using student data collected with passive sensing techniques, an area that has not been explored in existing literature. There were three main goals in this study. The first was to find the potential of behavioural features extracted from passive sensing data to distinguish between social and emotional loneliness among college students. The second aim was to determine whether these behavioural features can accurately predict and classify these loneliness types. The final aim was to identify the most important behavioural features that predictive models use to differentiate between two loneliness types. We used the two-sided Mann–Whitney U test to compare the distributions of the behavioural features between socially and emotionally lonely groups. The goal of this test was to determine if there were statistically significant differences in the behaviours of socially lonely and emotionally lonely groups. Then, we used machine learning classification models to identify the predictive power of identified behavioural features. Finally, we applied the importance analysis of SHAP features to find the most important features that were used by machine learning models for classification.

The results showed significant differences between the behavioural profiles of socially and emotionally lonely groups. Socially lonely individuals show less location variance, fewer location transitions, shorter phone usage duration, and lower Bluetooth scans than emotionally lonely individuals. Additionally, the socially lonely group had reduced physical activity and shorter awake time and longer sleep duration. Our predictive modelling found that the identified digital behavioural markers can effectively classify loneliness categories. Machine learning models with different algorithms significantly outperformed baseline approaches. XGBoost emerged as the top performer with an accuracy of 78.48%. SVM and Random Forest also demonstrated strong classification capabilities. Our SHAP analysis found that phone usage and location-based features were the most influential predictors of loneliness categories. For both XGBoost and Random Forest models, features like maximum phone usage duration, location variance, and location entropy were particularly important. The phone unlock patterns, and the number of significant places also played an important role.

Research shows that loneliness and depression often co-occur and can have a complex relationship with each other. Based on this, in Chapter 7, we worked on understanding the relationship between these two conditions using passive sensing data, which is a relatively unexplored approach in this context. This chapter aimed to answer several key questions regarding the behavioural aspects of loneliness and depression in college students. We looked into the different behavioural patterns linked to each condition, how they might be related and how they might act as mediators, and how well behavioural features combined with standard self-report scales could accurately detect loneliness and depression. We used two methods to understand how loneliness and depression show up in behaviours. First was logistic regression, which helped us identify specific behavioural differences between lonely/depressed students and others. Association rule mining (ARM) revealed complex patterns of behaviours linked to these conditions. This combination of methods provides a richer picture of how loneliness and depression manifest behaviourally in college students. Then, we used mediation analysis to examine the interplay between loneliness and depression. This analysis helped us identify specific behavioural features that mediate the relationship between these conditions which provided us findings into the underlying mechanisms of their connection. And finally, we trained multiple machine learning models with logistic regression, k-nearest neighbours, support vector machine (SVM), random forest, gradient boosting, and extreme gradient boosting to assess the power of combining behavioural features with traditional self-report scales to classify lonely and depressed students separately.

Our analysis found important behavioural differences between lonely and depressed students. Loneliness was associated with increased phone use (especially in the evening), more locations visited and reduced incoming calls. On the other hand, depressed people show fewer locations visited, less physical activity, shorter sleep duration, fewer conversations, and fewer Bluetooth scans. Association rule mining further highlighted complex combinations of these behaviours as predictive of loneliness and depression. Our mediation analysis also found significant links between loneliness and depression. ML models show the potential to classify lonely vs not lonely and depressed vs not depressed accurately. Across both conditions, the XGBoost model consistently outperformed other classifiers.

This thesis has presented a series of feasibility studies that explored various aspects of loneliness detection using the two pre-collected passive sensing datasets. Collectively, these studies show the significant potential of passive sensing-based methods combined with machine learning for understanding and detecting loneliness from both technical and theoretical perspectives. This thesis work is a stepping stone towards clinical evaluation and the real-world deployment of these methods.

8.2 Limitations and Future Work

Our work can be improved and extended in many ways. Throughout this thesis, we have addressed specific limitations and future directions within each relevant chapter. Here, we synthesize and expand upon these points to provide a comprehensive overview of potential improvements, discussions and extensions to our work. Some key points are given below, which build upon the chapter-specific discussions and consider the broader implications of our research.

8.2.1 Generalisability

AI-driven approaches to mental health detection using passive sensing data have a limitation of generalizability across diverse populations. Similar to findings in recent research on depression risk prediction, these models often have an algorithmic bias that leads to performing less accurately in varied demographic and socioeconomic subgroups [ASM+24]. This limitation is equally applicable to our work on loneliness detection.

Our models, which are trained on student populations, do not account for how loneliness manifests across different demographic and socioeconomic subgroups. Behavioural patterns that indicate loneliness in one subgroup may have different implications in another. For instance, increased mobility might indicate loneliness in one demographic group but not in another. Our approach of using clustering to identify behavioural subgroups and develop group-specific loneliness detection models (Chapter 4 and Chapter 5) presents a step towards addressing this heterogeneity in loneliness manifestation across different populations. However, it is important to note that it still has some limitations. Even within identified behavioural subgroups, meaningful variations may still exist that our models do not capture. Our grouping is based primarily on behavioural data. It does not account for other relevant factors that could influence loneliness manifestation. As we discussed in Chapter 7 as well, individuals in lower socioeconomic groups showed a positive association between mobility and depression risk, while higher socioeconomic groups showed the opposite [ASM+24]. This contradicts prior research, which found that mobility generally decreases depression risk [SZK+15]. Although this finding is about depression, the nature of this limitation is equally applicable to loneliness as well. Similarly, phone usage patterns associated with depression risk also differ significantly between age groups. Research shows that higher daytime phone use in younger individuals (18-25) is linked to lower depression risk, maybe due to its role in social connection and entertainment. However, this relationship is the opposite in older adults (65-74), where higher daytime phone use is associated with increased depression risk [ASM+24]. Again, the nature of such limitations is equally applicable to loneliness. As we have shown, the groups evolve over time, and behavioural subgroups can change. This dynamic nature presents ongoing challenges for long-term model accuracy. While our approach improves reliability within student samples, its effectiveness for significantly different populations (e.g., non-students or students from different cultural backgrounds) remains uncertain. Although our subgroup approach reduces some biases, it may inadvertently introduce others if certain demographic groups are overrepresented in particular behavioural clusters. This also raises important points about the fairness of these models.

We performed some experiments related to this issue by using the StudentLife dataset and our own collected UCC dataset with the younger population for a pilot study. The findings from these experiments further highlight these generalizability challenges. Despite both the StudentLife and UCC datasets being collected from younger populations, we observed significant performance drops when applying models trained on one dataset to the other. Specifically, the loneliness detection model trained on the StudentLife dataset performed poorly when tested on the UCC dataset and vice versa. However, performance improved when we trained and tested a model using only the UCC dataset and the StudentLife dataset.

While these experiments highlight significant challenges in generalizing loneliness detection models, our research also shows the potential benefits of personalization (Chapter 4 and Chapter 5). This raises important questions about the trade-off between generalization vs. personalization. On the one hand, generalizable models can be used for broader, diverse populations. However, as our findings show, these models may perform poorly when applied to populations different from their training data and when increasing population size. On the other hand, personalized models showed improved reliability within their target populations. However, the development of such personalized models for every possible subgroup could be resource-intensive and reduce the scalability of loneliness detection systems. This trade-off between generalization vs. personalization opens up an important question for future research: To what extent can we develop models that are both generalizable across populations and sensitive to individual or group-specific variations in loneliness manifestation?

These findings show both the limitations and potential of passive sensing-based models for loneliness detection. While these models show promise when personalized, they are not yet robust enough for diverse populations. Future work needs to include a broader range of data, including detailed demographic information, socioeconomic indicators, and other relevant contextual factors. This would help for more nuanced and comprehensive understanding that could potentially lead to more robust and accurate loneliness detection models.

8.2.2 Longitudinal Studies

Another limitation of this thesis is the short duration of the datasets used. The datasets we used were collected over a 10-week period, which is insufficient for capturing the gradual and nuanced behavioural changes associated with loneliness. Loneliness is typically a gradual evolving experience rather than an abrupt change in emotional state. So, the behavioural manifestations of loneliness may take considerable time to become apparent or may fluctuate subtly over longer periods. The 10-week timeframe may miss these nuanced, long-term patterns of change.

Future studies should aim to collect long-term data. This would help researchers to observe and analyze the subtle, progressive changes in behaviour that may indicate the onset or intensification of loneliness. These long-term periods could find cyclical patterns in loneliness and associated behaviours, which might be linked to academic cycles, seasons, or life stages. Moreover, adding time series analysis in future studies could capture moments of change and provide a deeper understanding of the temporal dynamics of loneliness and behavioural changes. This approach would help detect critical periods where interventions might be most effective and could lead to more precise and timely loneliness detection and intervention strategies.

8.2.3 Limitations of Binary Classification

We have mainly used binary classification for loneliness throughout this thesis. While this approach simplified our analysis, it does not fully capture loneliness's complex, continuous nature. The reason for using binary classification was due to small datasets with only a 10-week period. As we have discussed above, loneliness typically changes slowly, so trying to categorize it into multiple levels or treat it as a continuous variable would have been tricky with such limited data. This could have led to reduced statistical power, increased risk of overfitting, and potential reliability issues in distinguishing between fine-grained loneliness levels over such a short period. By using binary classification, we could more clearly see the behavioural patterns linked to lonely versus not lonely. Moreover, binary outcomes tend to give more stable and easy-to-interpret results in ML, especially with smaller datasets. However, we recognize this approach might have missed some of the subtle variations in how loneliness can intensify or change over time.

Another important limitation associated with this binary classification approach is the selection of a threshold score from the UCLA scale to divide participants into lonely/not lonely categories. This choice of threshold directly affects the distribution of students across the lonely and not lonely categories. A higher or lower threshold could lead to different categorizations that can potentially classify more or fewer participants as lonely. For example, some students who are close to the cutoff score may experience loneliness but are not classified as such due to the arbitrary selection of this threshold. This has important implications for the model's performance, as small variations in the threshold can change the balance between the two classes (lonely/not lonely), thus affecting the overall classification accuracy, sensitivity, and precision. The binary classification using this threshold may oversimplify loneliness. Students near the cutoff, called borderline cases, might show behaviours that are more complex than the "lonely" or "not lonely" binary classification. This means the model could miss subtle changes in their loneliness, especially for those whose scores fluctuate around the threshold. As a result, important behaviours from students just below the threshold may be ignored or misclassified. While we did not conduct a sensitivity analysis in this thesis, future work could include this approach to better understand how the model handles borderline participants. Exploring techniques such as fuzzy classification or applying threshold tuning based on behavioural patterns over time could reduce misclassification. Moreover, as we discussed above, future studies with longitudinal data collection could play a key role here as well. This could provide a more detailed analysis of how these borderline cases fluctuate that can help to develop more robust models that can accommodate the nuanced changes in loneliness levels.

8.2.4 Limitations of Cross-Validation Approaches

Throughout the thesis, we have used different cross-validation approaches to evaluate model performances. In Chapters 4 and 5, we used stratified 10-fold cross-validation which is a widely used technique that divides the data into 10 balanced folds to ensure each fold has a relatively even distribution of lonely and non-lonely cases. This method provided stable and balanced performance metrics. However, an inherent limitation of 10-fold cross-validation in our context is the potential for data leakage. Since each student's data could appear in both the training and test folds across multiple iterations, there is a risk of the model indirectly learning individual behavioural patterns through repeated exposure in training. This could lead to inflated performance metrics, as the model might benefit from familiarity with each student's behaviour across folds. In Chapters 6 and 7, we used LOSO cross-validation to mitigate this risk by fully isolating each student's data during testing, which provides a stricter evaluation but also potentially results in lower metrics due to this isolation.

Future work should consider conducting LOSO-based analyses in Chapters 4 and 5 to achieve a consistent and fair evaluation across all chapters. This would help to ensure that performance improvements we observed in group-specific or generalized models are not influenced by unintentional data leakage and that the models can generalize reliably to new students.

8.2.5 Algorithmic Improvements

Throughout this thesis, we have consistently used traditional ML algorithms for loneliness detection. Our primary focus was to explore and understand the behavioural manifestations of loneliness, which are very easy to do with traditional ML models as they are easy to interpret. The relatively small datasets that are common in mental health research are better suited to these methods and are less prone to overfitting than deep learning models. Another reason is that these algorithms have been widely used in mental health research which provides a solid foundation for comparison and validation of our findings within the existing literature.

While these approaches served our research goals effectively, future work should explore the potential of deep learning in loneliness detection. However, the deployment of deep learning models requires substantially larger datasets to train effectively without overfitting, and usually deep learning-based models are harder to interpret than traditional ML models. As we discussed above, larger datasets over longer periods would address the limitations in capturing gradual behavioural changes and provide the volume of data necessary to train and validate deep learning models effectively. Another option could be to investigate the use of domain adaptation and transfer learning techniques. This would involve using models pre-trained on larger, more diverse datasets (e.g., datasets related to mental health, well-being, or even general behaviour) and then fine-tuning them specifically for the task of loneliness detection.

8.2.6 Self-Reported Bias

Another limitation in our methodology is the reliance on self-reported loneliness surveys to label passive sensing data. The subjective nature of such reporting may lead to discrepancies between reported feelings and actual emotional states due to factors like response bias, mood variation, limited self-awareness, and differing interpretations of survey questions. Cultural factors and temporal mismatches between survey completion and data collection periods can further complicate this issue. These factors potentially affect the quality of labels used to train our machine-learning models which could impact their accuracy and reliability. Future research should address these limitations by incorporating multiple assessment methods, increasing the frequency of surveys, developing more objective measures and conducting validation studies.

8.2.7 Establishing Causality

It is important to emphasize that while our research has identified significant associations between various behavioural features and loneliness, these findings do not imply causality. The relationships we have found between digital behavioural markers and loneliness are correlational in nature, and we must be cautious in interpreting these results. We cannot conclude that these behaviours directly cause loneliness or vice versa. The relationship between behaviours and loneliness is complex and bidirectional and could be influenced by multiple factors that our current methodology may not fully capture. We only have a partial view of a person's life through sensor data. There is a lot more going on in a person's life that could play a significant role in relation to loneliness. For instance, passive sensing might detect low activity levels or disrupted sleep patterns, which are often associated with loneliness. However, these same patterns could be indicative of various other conditions or circumstances. For example, if an individual shows reduced physical activity and irregular sleep patterns, although these behaviours align with potential loneliness indicators, they could also be symptoms of physical illness, work-related stress, or even a period of intense focus on a project. Without additional contextual information, these models risk misinterpreting these patterns and potentially misclassifying an individual's mental state.

Similarly for causality, while we found that increased phone usage correlates with higher levels of loneliness, we cannot determine whether excessive phone use leads to loneliness, if lonely individuals tend to use their phones more, or if other underlying factors influence both. Broader spectrum data would be needed to get a broader picture that could lead to more robust and accurate detection models and to establish causality. Additionally, these behavioural patterns should be monitored over extended periods along with contextual data to distinguish between temporary fluctuations in behavioural indicators and persistent trends that could truly indicate loneliness.

8.2.8 Interdisciplinary Approaches

Future work also needs to incorporate interdisciplinary approaches for the validation of these passive sensing based systems within clinical settings. This collaboration with clinical psychologists can give us a more comprehensive understanding of loneliness as a multifaceted phenomenon and its various behavioural indicators. This would also help in validation and reliability of using such passive sensing based systems in real world scenarios. By integrating information from different disciplines, researchers can develop more accurate and reliable models that align closely with clinical definitions of loneliness and its associated behaviours. This collaboration would help in ethical considerations in a more granular way that would improve the overall impact and acceptability of such systems in wider populations.

8.2.9 Timely Interventions

The ultimate goal is to detect loneliness early, which would result in timely interventions before loneliness leads to more severe mental and physical health issues. Future research should focus on designing and testing intervention systems that can be automatically triggered by signals detected from the sensing systems. Interventions could be of several types, like app messages offering a bit of encouragement or reminding someone to reach out to a friend. For more serious cases, alerts could even go to a therapist. The key is linking those interventions directly to what the detection systems are picking up on, which would create a support system that doesn't just monitor but actually helps.

8.3 Final Comment

Building upon the foundational work of previous researchers in the field, this thesis further pushes the boundaries of understanding and detecting loneliness through passive sensing techniques. Future directions, including expanded datasets with different populations, longitudinal studies, investigating generalisation vs. personalisation and cross-disciplinary collaboration and validation, can help to refine these methods and unlock their full potential. I am optimistic that continued research and collaboration will result in more robust, nuanced tools for detecting both loneliness and other related mental health conditions. These advancements have the power to improve our understanding of these conditions, early detection and more importantly to facilitate timely, targeted interventions that could transform the lives of people worldwide. While much work remains, the findings and potential presented here give me hope that we can create a future where loneliness is detected at early stages, leading to timely intervention, and this way, its harmful effects could be minimized. Researchers, clinicians, and technologists can work together to translate these findings into real, tangible detection and support systems, which would be a significant step toward combating this pervasive challenge of our time.

Acronyms

ARM Association Rule Mining.

EL Emotional Loneliness.

EMA Ecological Momentary Assessment.

EPS epsilon.

EVR Explained Variance Ratio.

GLOBEM Generalization of LOngitudinal BEhavior Modeling.

GPS Global Positioning System.

IDBSCAN Incremental Density-based Spatial Clustering of Applications with Noise.

IRB Institutional Review Board.

MinPts Minimum Points.

ML Machine Learning.

OPTICS Ordering Points To Identify the Clustering Structure.

OR Odds Ratio.

PAM Partition Around Medoids.

PANAS Positive and Negative Affect Schedule.

PCA Principal Component Analysis.

PHQ Patient Health Questionnaire.

PRISMA Preferred Reporting Items for Systematic Reviews and Meta-Analyses.

RAPIDS Reproducible Analysis Pipeline for Data Streams.

SL Social Loneliness.

SMOTE Synthetic Minority Oversampling Method.

SVM Support Vector Machine.

UCLA University of California, Los Angeles.

XGBoost Extreme Gradient Boosting.

References

[AA22] Md Sabbir Ahmed and Nova Ahmed. Less is more: Leveraging digital behavioral markers for real-time identification of loneliness in resource-limited settings. In *International Conference on Pervasive Computing Technologies for Healthcare*, pages 460–476. Springer, 2022.

[AAT+18] Aljohara A Alhassan, Ethar M Alqadhib, Nada W Taha, Raneem A Alahmari, Mahmoud Salam, and Adel F Almutairi. The relationship between addiction to smartphone usage and depression among adults: a cross sectional study. *BMC psychiatry*, 18:1–8, 2018.

[AB13] Kirstie N Anderson and Andrew J Bradley. Sleep disturbance in mental health problems and neurodegenerative disease. *Nature and science of sleep*, pages 61–75, 2013.

[ABKS99] Mihael Ankerst, Markus M Breunig, Hans-Peter Kriegel, and Jörg Sander. Optics: Ordering points to identify the clustering structure. *ACM Sigmod record*, 28(2):49–60, 1999.

[ADR+16] Johanna Austin, Hiroko H Dodge, Thomas Riley, Peter G Jacobs, Stephen Thielke, and Jeffrey Kaye. A smart-home system to unobtrusively and continuously assess loneliness in older adults. *IEEE journal of translational engineering in health and medicine*, 4:1–11, 2016.

[AIS93] Rakesh Agrawal, Tomasz Imieliński, and Arun Swami. Mining association rules between sets of items in large databases. In *Proceedings of the 1993 ACM SIGMOD international conference on Management of data*, pages 207–216, 1993.

[AL06] Bradley M Appelhans and Linda J Luecken. Heart rate variability as an index of regulated emotional responding. *Review of general psychology*, 10(3):229–240, 2006.

[AO05] Hilary Arksey and Lisa O'Malley. Scoping studies: towards a methodological framework. *International journal of social research methodology*, 8(1):19–32, 2005.

[AOASAR+23] Omar Al Omari, Sulaiman Al Sabei, Omar Al Rawajfah, Loai Abu Sharour, Iman Al-Hashmi, Mohammad Al Qadire, and Atika Khalaf. Prevalence and predictors of loneliness among youth during the time of covid-19: a multinational study. *Journal of the American Psychiatric Nurses Association*, 29(3):204–214, 2023.

[Ars21] Gökmen Arslan. Loneliness, college belongingness, subjective vitality, and psychological adjustment during coronavirus pandemic: Development of the college belongingness questionnaire. *Journal of Positive School Psychology*, 5(1):17–31, 2021.

[ASM+24] Daniel A Adler, Caitlin A Stamatis, Jonah Meyerhoff, David C Mohr, Fei Wang, Gabriel J Aranovich, Srijan Sen, and Tanzeem Choudhury. Measuring algorithmic bias to analyze the reliability of ai tools that predict depression risk using smartphone sensed-behavioral data. *npj Mental Health Research*, 3(1):17, 2024.

[BBDC+21] Julia Baarck, Alexandra BALAHUR-DOBRESCU, Laura Giulia CASSIO, Beatrice D'HOMBRES, Zsuzsanna Pásztor, Guido Tintori, et al. Loneliness in the eu. insights from surveys and online media data. 2021.

[BCC+21] Gloria Bellini, Marco Cipriano, Sara Comai, Nicola De Angeli, Jacopo Pio Gargano, Matteo Gianella, Gianluca Goi, Giovanni Ingrao, Andrea Masciadri, Gabriele Rossi, et al. Understanding social behaviour in a health-care facility from localization data: a case study. *Sensors*, 21(6):2147, 2021.

[BG88] Robert A Bell and M Christina Gonzalez. Loneliness, negative life events, and the provisions of social relationships. *Communication Quarterly*, 36(1):1–15, 1988.

[BK20] Christopher R Beam and Alice J Kim. Psychological sequelae of social isolation and loneliness might be a larger problem in young adults than older adults. *Psychological Trauma: Theory, Research, Practice, and Policy*, 12(S1):S58, 2020.

[BL17] Roy F Baumeister and Mark R Leary. The need to belong: Desire for interpersonal attachments as a fundamental human motivation. *Interpersonal development*, pages 57–89, 2017.

[BNW+18] Tjeerd W Boonstra, Jennifer Nicholas, Quincy JJ Wong, Frances Shaw, Samuel Townsend, and Helen Christensen. 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):e10131, 2018.

[BPR+08] África Borges, Pedro Prieto, Giacinto Ricchetti, Carmen Hernández-Jorge, and Elena Rodríguez-Naveiras. Validación cruzada de la factorización del test ucla de soledad. *Psicothema*, 20(4):924–927, 2008.

[BSF20] Feifei Bu, Andrew Steptoe, and Daisy Fancourt. Who is lonely in lockdown? cross-cohort analyses of predictors of loneliness before and during the covid-19 pandemic. *Public health*, 186:31–34, 2020.

[Bur95] Jerry M Burger. Individual differences in preference for solitude. *Journal of Research in Personality*, 29(1):85–108, 1995.

[BWA+23] Amelia Blankenau, Amy Wax, Lena Auerbach, Zachary D Schuman, and Andrea Hopmeyer. Queer peer crowds on campus: Lgbt crowd affiliation as a critical correlate of college students' loneliness, academic well-being, & stress. *Journal of Homosexuality*, 70(7):1411–1439, 2023.

[BZSW+15] Dror Ben-Zeev, Emily A Scherer, Rui Wang, Haiyi Xie, and Andrew T Campbell. Next-generation psychiatric assessment: Using smartphone sensors to monitor behavior and mental health. *Psychiatric rehabilitation journal*, 38(3):218, 2015.

[Cap06] Scott E Caplan. Relations among loneliness, social anxiety, and problematic internet use. *CyberPsychology & behavior*, 10(2):234–242, 2006.

[CBHK02] Nitesh V Chawla, Kevin W Bowyer, Lawrence O Hall, and W Philip Kegelmeyer. Smote: synthetic minority over-sampling technique. *Journal of artificial intelligence research*, 16:321–357, 2002.

[CBY+20] Xinli Chi, Benjamin Becker, Qian Yu, Peter Willeit, Can Jiao, Liuyue Huang, M Mahhub Hossain, Igor Grabovac, Albert Yeung, Jingyuan Lin, et al. Prevalence and psychosocial correlates of mental health outcomes among chinese college students during the coronavirus disease (covid-19) pandemic. *Frontiers in psychiatry*, 11:803, 2020.

[CC08] A Aykut Ceyhan and Esra Ceyhan. Loneliness, depression, and computer self-efficacy as predictors of problematic internet use. *CyberPsychology & Behavior*, 11(6):699–701, 2008.

[CEB+00] John T Cacioppo, John M Ernst, Mary H Burleson, Martha K McClintock, William B Malarkey, Louise C Hawkley, Ray B Kowalewski, Alisa Paulsen, J Allan Hobson, Kenneth Hugdahl, et al. Lonely traits and concomitant physiological processes: The macarthur social neuroscience studies. *International Journal of Psychophysiology*, 35(2-3):143–154, 2000.

[CH03] John T Cacioppo and Louise C Hawkley. Social isolation and health, with an emphasis on underlying mechanisms. *Perspectives in biology and medicine*, 46(3):S39–S52, 2003.

[CH09] John T Cacioppo and Louise C Hawkley. Perceived social isolation and cognition. *Trends in cognitive sciences*, 13(10):447–454, 2009.

[CH18] Victor P Cornet and Richard J Holden. Systematic review of smartphone-based passive sensing for health and wellbeing. *Journal of biomedical informatics*, 77:120–132, 2018.

[CHC+02] John T Cacioppo, Louise C Hawkley, L Elizabeth Crawford, John M Ernst, Mary H Burleson, Ray B Kowalewski, William B Malarkey, Eve Van Cauter, and Gary G Berntson. Loneliness and health: Potential mechanisms. *Psychosomatic medicine*, 64(3):407–417, 2002.

[CHE+06] John T Cacioppo, Louise C Hawkley, John M Ernst, Mary Burleson, Gary G Berntson, Bita Nouriani, and David Spiegel. Loneliness within a nomological net: An evolutionary perspective. *Journal of research in personality*, 40(6):1054–1085, 2006.

[CHT10] John T Cacioppo, Louise C Hawkley, and Ronald A Thisted. Perceived social isolation makes me sad: 5-year cross-lagged analyses of loneliness and depressive symptomatology in the chicago health, aging, and social relations study. *Psychology and aging*, 25(2):453, 2010.

[CKC+15] Sam-Wook Choi, Dai-Jin Kim, Jung-Seok Choi, Heejune Ahn, Eun-Jeung Choi, Won-Young Song, Seohee Kim, and Hyunchul Youn. Comparison of risk and protective factors associated with smartphone addiction and internet addiction. *Journal of behavioral addictions*, 4(4):308–314, 2015.

[Coh13] Jacob Cohen. *Statistical power analysis for the behavioral sciences*. routledge, 2013.

[CP08] John T Cacioppo and William Patrick. *Loneliness: Human nature and the need for social connection*. WW Norton & Company, 2008.

[CRC+19] Angelo Costa, Jaime A Rincon, Carlos Carrascosa, Vicente Julian, and Paulo Novais. Emotions detection on an ambient intelligent system using wearable devices. *Future Generation Computer Systems*, 92:479–489, 2019.

[DALW+22] Valeria De Angel, Serena Lewis, Katie White, Carolin Oetzmann, Daniel Leightley, Emanuela Oprea, Grace Lavelle, Faith Matcham, Alice Pace, David C Mohr, et al. Digital health tools for the passive monitoring of depression: a systematic review of methods. *NPJ digital medicine*, 5(1):3, 2022.

[Dat22] K. Viraj Datt. Incremental machine learning for streaming data with river - part 3: Classification algorithms. `https://kvirajdatt.medium.com/incremental-machine-learning-for-streaming-data-with-river-part-3-classification-algorithms-315c145dcb92`, 2022.

[Dat24] DataCamp. What is incremental learning? `https://www.datacamp.com/blog/what-is-incremental-learning`, 2024.

[dJG89] Jenny de Jong-Gierveld. Personal relationships, social support, and loneliness. *Journal of Social and Personal Relationships*, 6(2):197–221, 1989.

[DJGK85] Jenny De Jong-Gierveld and Frans Kamphuls. The development of a rasch-type loneliness scale. *Applied psychological measurement*, 9(3):289–299, 1985.

[DJGVT99] Jenny De Jong Gierveld and Theo Van Tilburg. Manual of the loneliness scale 1999. *Department of Social Research Methodology, Vrije Universiteit Amsterdam, Amsterdam (updated version 18.01. 02)*, 1999.

[DJIHK18] Katharina Diehl, Charlotte Jansen, Kamila Ishchanova, and Jennifer Hilger-Kolb. Loneliness at universities: determinants of emotional and social loneliness among students. *International journal of environmental research and public health*, 15(9):1865, 2018.

[DVC+19] Afsaneh Doryab, Daniella K Villalba, Prerna Chikersal, Janine M Dutcher, Michael Tumminia, Xinwen Liu, Sheldon Cohen, Kasey Creswell, Jennifer Mankoff, John D Creswell, et al. Identifying behavioral phenotypes of loneliness and social isolation with passive sensing: statistical analysis, data mining and machine learning of smartphone and fitbit data. *JMIR mHealth and uHealth*, 7(7):e13209, 2019.

[EA19] Dina Elreedy and Amir F Atiya. A comprehensive analysis of synthetic minority oversampling technique (smote) for handling class imbalance. *Information Sciences*, 505:32–64, 2019.

[EDKN+16] Asli Enez Darcin, Samet Kose, Cemal Onur Noyan, Serdar Nurmedov, Onat Yılmaz, and Nesrin Dilbaz. Smartphone addiction and its relationship with social anxiety and loneliness. *Behaviour & Information Technology*, 35(7):520–525, 2016.

[EGH+14] William Enck, Peter Gilbert, Seungyeop Han, Vasant Tendulkar, Byung-Gon Chun, Landon P Cox, Jaeyeon Jung, Patrick McDaniel, and Anmol N Sheth. Taintdroid: an information-flow tracking system for realtime privacy monitoring on smartphones. *ACM Transactions on Computer Systems (TOCS)*, 32(2):1–29, 2014.

[EKS+98] Martin Ester, Hans-Peter Kriegel, Jorg Sander, Michael Wimmer, and Xiaowei Xu. Incremental clustering for mining in a data ware housing. *University of Munich Oettingenstr*, 67, 1998.

[EL23] Timon Elmer and Gerine Lodder. Modeling social interaction dynamics measured with smartphone sensors: An ambulatory assessment study on social interactions and loneliness. *Journal of Social and Personal Relationships*, 40(2):654–669, 2023.

[ETW+18] Jon D Elhai, Mojisola F Tiamiyu, Justin W Weeks, Jason C Levine, Kristina J Picard, and Brian J Hall. Depression and emotion regulation predict objective smartphone use measured over one week. *Personality and individual differences*, 133:21–28, 2018.

[FGP08] Katayoun Farrahi and Daniel Gatica-Perez. What did you do today? discovering daily routines from large-scale mobile data. In *Proceedings of the 16th ACM international conference on Multimedia*, pages 849–852, 2008.

[FK07] Tineke Fokkema and Kees Knipscheer. Escape loneliness by going digital: A quantitative and qualitative evaluation of a dutch experiment in using ect to overcome loneliness among older adults, 2007.

[FMG+20] Daniel Fulford, Jasmine Mote, Rachel Gonzalez, Samuel Abplanalp, Yuting Zhang, Jarrod Luckenbaugh, Jukka-Pekka Onnela, Carlos Busso, and David E Gard. Smartphone sensing of social interactions in people with and without schizophrenia. *Journal of Psychiatric Research*, 2020.

[FMG+21] Daniel Fulford, Jasmine Mote, Rachel Gonzalez, Samuel Abplanalp, Yuting Zhang, Jarrod Luckenbaugh, Jukka-Pekka Onnela, Carlos Busso, and David E Gard. Smartphone sensing of social interactions in people with and without schizophrenia. *Journal of Psychiatric Research*, 137:613–620, 2021.

[Fra24] AWARE Framework. Aware sensors details. `https://awareframework.com/sensors/`, 2024.

[FSRB22] José Marcelo Fernandes, Jorge Sá Silva, André Rodrigues, and Fernando Boavida. A survey of approaches to unobtrusive sensing of humans. *ACM Computing Surveys (CSUR)*, 55(2):1–28, 2022.

[GCRN+18] Enrique Garcia-Ceja, Michael Riegler, Tine Nordgreen, Petter Jakobsen, Ketil J Oedegaard, and Jim Tørresen. Mental health monitoring with multimodal sensing and machine learning: A survey. *Pervasive and Mobile Computing*, 51:1–26, 2018.

[GLZ14] Yusong Gao, He Li, and Tingshao Zhu. Predicting subjective well-being by smartphone usage behaviors. In *HEALTHINF*, pages 317–322, 2014.

[GLZ+16] Yusong Gao, Ang Li, Tingshao Zhu, Xiaoqian Liu, and Xingyun Liu. How smartphone usage correlates with social anxiety and loneliness. *PeerJ*, 4:e2197, 2016.

[GM11] Scott A Golder and Michael W Macy. Diurnal and seasonal mood vary with work, sleep, and daylength across diverse cultures. *Science*, 333(6051):1878–1881, 2011.

[GNH+17] Fei Gu, Jianwei Niu, Zhenxue He, Xin Jin, and Joel JPC Rodrigues. Smartbuddy: an integrated mobile sensing and detecting system for family activities. In *GLOBECOM 2017-2017 IEEE Global Communications Conference*, pages 1–7. IEEE, 2017.

[GT06] Jenny De Jong Gierveld and Theo Van Tilburg. A 6-item scale for overall, emotional, and social loneliness: Confirmatory tests on survey data. *Research on aging*, 28(5):582–598, 2006.

[GTT17] Nadee Goonawardene, XiaoPing Toh, and Hwee-Pink Tan. Sensor-driven detection of social isolation in community-dwelling elderly. In *International Conference on Human Aspects of IT for the Aged Population*, pages 378–392. Springer, 2017.

[GVT10] Jenny De Jong Gierveld and Theo Van Tilburg. The de jong gierveld short scales for emotional and social loneliness: tested on data from 7 countries in the un generations and gender surveys. *European journal of ageing*, 7(2):121–130, 2010.

[GWR+20] Sarah C Griffin, Allison B Williams, Scott G Ravyts, Samantha N Mladen, and Bruce D Rybarczyk. Loneliness and sleep: A systematic review and meta-analysis. *Health psychology open*, 7(1):2055102920913235, 2020.

[GZPA23] Beni Gómez-Zúñiga, Modesta Pousada, and Manuel Armayones. Loneliness and disability: A systematic review of loneliness conceptualization and intervention strategies. *Frontiers in Psychology*, 13:1040651, 2023.

[HBC05] Louise C Hawkley, Michael W Browne, and John T Cacioppo. How can i connect with thee? let me count the ways. *Psychological Science*, 16(10):798–804, 2005.

[HCH+22] Shaun Hayes, Molly Carlyle, S Alexander Haslam, Catherine Haslam, and Genevieve Dingle. Exploring links between social identity, emotion regulation, and loneliness in those with and without a history of mental illness. *British Journal of Clinical Psychology*, 61(3):701–734, 2022.

[HD87] Ron D Hays and M Robin DiMatteo. A short-form measure of loneliness. *Journal of personality assessment*, 51(1):69–81, 1987.

[HG06] Liesl M Heinrich and Eleonora Gullone. The clinical significance of loneliness: A literature review. *Clinical psychology review*, 26(6):695–718, 2006.

[HG08] Kyle Heath and Leonidas Guibas. Multi-person tracking from sparse 3d trajectories in a camera sensor network. In *2008 Second ACM/IEEE International Conference on Distributed Smart Cameras*, pages 1–9. IEEE, 2008.

[HLHS+22] Emily Hards, Maria Elizabeth Loades, Nina Higson-Sweeney, Roz Shafran, Teona Serafimova, Amberley Brigden, Shirley Reynolds, Esther Crawley, Eleanor Chatburn, Catherine Linney, et al. Loneliness and mental health in children and adolescents with pre-existing mental health problems: A rapid systematic review. *British Journal of Clinical Psychology*, 61(2):313–334, 2022.

[HLMA19] J Henriksen, ER Larsen, C Mattisson, and NW Andersson. Loneliness, health and mortality. epidemiol. psychiatr. sci. 28 (2): 234–239, 2019.

[HMY22] Md Mehedi Hassan, Swarnali Mollick, and Farhana Yasmin. An unsupervised cluster-based feature grouping model for early diabetes detection. *Healthcare Analytics*, 2:100112, 2022.

[HÖK+22] Jessica Hemberg, Lillemor Östman, Yulia Korzhina, Henrik Groundstroem, Lisbet Nyström, and Pia Nyman-Kurkiala. Loneliness as experienced by adolescents and young adults: An explorative qualitative study. *International Journal of Adolescence and Youth*, 27(1):362–384, 2022.

[HPBGP17] Rui Hu, Hieu Pham, Philipp Buluschek, and Daniel Gatica-Perez. Elderly people living alone: Detecting home visits with ambient and wearable sensing. In *Proceedings of the 2nd International Workshop on Multimedia for Personal Health and Health Care*, pages 85–88, 2017.

[HSSW20] Louise C Hawkley, Andrew Steptoe, L Philip Schumm, and Kristen Wroblewski. Comparing loneliness in england and the united states, 2014–2016: Differential item functioning and risk factor prevalence and impact. *Social Science & Medicine*, 265:113467, 2020.

[HTC09] Louise C Hawkley, Ronald A Thisted, and John T Cacioppo. Loneliness predicts reduced physical activity: cross-sectional & longitudinal analyses. *Health psychology*, 28(3):354, 2009.

[HTL17] Sinh Huynh, Hwee-Pink Tan, and Youngki Lee. Towards unobtrusive mental well-being monitoring for independent-living elderly. In *Proceedings of the 4th International on Workshop on Physical Analytics*, pages 1–6, 2017.

[IBY21] Kevser Isik, Ceyda Başoğul, and Hilal Yildirim. The relationship between perceived loneliness and depression in the elderly and influencing factors. *Perspectives in psychiatric care*, 57(1), 2021.

[JAL+23] Salar Jafarlou, Iman Azimi, Jocelyn Lai, Yuning Wang, Sina Labbaf, Brenda Nguyen, Hana Qureshi, Christopher Marcotullio, Jessica L Borelli, Nikil D Dutt, et al. Objective monitoring of loneliness levels using smart devices: A multi-device approach for mental health applications. *medRxiv*, pages 2023–06, 2023.

[JC16] Ian T Jolliffe and Jorge Cadima. Principal component analysis: a review and recent developments. *Philosophical transactions of the royal society A: Mathematical, Physical and Engineering Sciences*, 374(2065):20150202, 2016.

[JC20] Nicholas C Jacobson and Yeon Joo Chung. Passive sensing of prediction of moment-to-moment depressed mood among undergraduates with clinical levels of depression sample using smartphones. *Sensors*, 20(12):3572, 2020.

[JLY+23] Xiayan Ji, Xian Li, Ahhyun Yuh, Claire Kendell, Amanda Watson, James Weimer, Hajime Nagahara, Teruo Higashino, Teruhiro Mizumoto, Viktor Erdélyi, et al. Short: Integrated sensing platform for detecting social isolation and loneliness in the elderly community. In *Proceedings of the 8th ACM/IEEE International Conference on Connected Health: Applications, Systems and Engineering Technologies*, pages 148–152, 2023.

[JN13] Kalpana D Joshi and PS Nalwade. Modified k-means for better initial cluster centres. *International Journal of Computer Science and Mobile Computing*, 2(7):219–223, 2013.

[Joi23] Joint Research Centre, European Commission. Eu loneliness survey. `https://joint-research-centre.ec.europa.eu/scientific-activities-z/loneliness/eu-loneliness-survey_en`, 2023.

[JSW20] Nicholas C Jacobson, Berta Summers, and Sabine Wilhelm. Digital biomarkers of social anxiety severity: digital phenotyping using passive smartphone sensors. *Journal of medical Internet research*, 22(5):e16875, 2020.

[JTB+20] Louis Jacob, Mark A Tully, Yvonne Barnett, Guillermo F Lopez-Sanchez, Laurie Butler, Felipe Schuch, Rubén López-Bueno, Daragh McDermott, Joseph Firth, Igor Grabovac, et al.

The relationship between physical activity and mental health in a sample of the uk public: A cross-sectional study during the implementation of covid-19 social distancing measures. *Mental health and physical activity*, 19:100345, 2020.

[KAGH20] Gülsen Kılınç, Rukuye Aylaz, Gülsen Güneş, and Pınar Harmancı. The relationship between depression and loneliness levels of the students at the faculty of health sciences and the factors affecting them. *Perspectives in Psychiatric Care*, 56(2):431–438, 2020.

[KCMG88] Robert G Knight, Barbara J Chisholm, Nigel V Marsh, and Hamish PD Godfrey. Some normative, reliability, and factor analytic data for the revised ucla loneliness scale. *Journal of Clinical Psychology*, 44(2):203–206, 1988.

[KKH+11] Lianne M Kurina, Kristen L Knutson, Louise C Hawkley, John T Cacioppo, Diane S Lauderdale, and Carole Ober. Loneliness is associated with sleep fragmentation in a communal society. *Sleep*, 34(11):1519–1526, 2011.

[KL19] Anissa Lestari Kadiyono and Annisa Nanda Liyani. Finding the gap: Student stress and student engagement of first-year undergraduate student from overseas. In *SU-AFBE 2018: Proceedings of the 1st Sampoerna University-AFBE International Conference, SU-AFBE 2018, 6-7 December 2018, Jakarta Indonesia*, page 80. European Alliance for Innovation, 2019.

[KSW01] Kurt Kroenke, Robert L Spitzer, and Janet BW Williams. The phq-9: validity of a brief depression severity measure. *Journal of general internal medicine*, 16(9):606–613, 2001.

[LAS+15] Bayard E Lyons, Daniel Austin, Adriana Seelye, Johanna Petersen, Jonathan Yeargers, Thomas Riley, Nicole Sharma, Nora Mattek, Katherine Wild, Hiroko Dodge, et al. Pervasive computing technologies to continuously assess alzheimer's disease progression and intervention efficacy. *Frontiers in aging neuroscience*, 7:102, 2015.

[LAT+09] Alessandro Liberati, Douglas G Altman, Jennifer Tetzlaff, Cynthia Mulrow, Peter C Gøtzsche, John PA Ioannidis, Mike

Clarke, Philip J Devereaux, Jos Kleijnen, and David Moher. The prisma statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. *Annals of internal medicine*, 151(4):W–65, 2009.

[LBD+21] Mira I Leese, John PK Bernstein, Katherine E Dorociak, Nora Mattek, Chao-Yi Wu, Zachary Beattie, Hiroko H Dodge, Jeffrey Kaye, and Adriana M Hughes. Older adults' daily activity and mood changes detected during the covid-19 pandemic using remote unobtrusive monitoring technologies. *Innovation in Aging*, 5(4):igab032, 2021.

[LCHS+20] Maria Elizabeth Loades, Eleanor Chatburn, Nina Higson-Sweeney, Shirley Reynolds, Roz Shafran, Amberly Brigden, Catherine Linney, Megan Niamh McManus, Catherine Borwick, and Esther Crawley. Rapid systematic review: the impact of social isolation and loneliness on the mental health of children and adolescents in the context of covid-19. *Journal of the American Academy of Child & Adolescent Psychiatry*, 59(11):1218–1239, 2020.

[LCO10] Danielle Levac, Heather Colquhoun, and Kelly K O'Brien. Scoping studies: advancing the methodology. *Implementation science*, 5(1):1–9, 2010.

[LDB19] David John Lenny, Tenzin Doleck, and Paul Bazelais. Self-determination, loneliness, fear of missing out, and academic performance. *Knowledge Management & E-Learning*, 11(4):485–496, 2019.

[LEC+20] Scott M Lundberg, Gabriel Erion, Hugh Chen, Alex DeGrave, Jordan M Prutkin, Bala Nair, Ronit Katz, Jonathan Himmelfarb, Nisha Bansal, and Su-In Lee. From local explanations to global understanding with explainable ai for trees. *Nature machine intelligence*, 2(1):56–67, 2020.

[LEV20] Michelle H Lim, Robert Eres, and Shradha Vasan. Understanding loneliness in the twenty-first century: an update on correlates, risk factors, and potential solutions. *Social psychiatry and psychiatric epidemiology*, 55:793–810, 2020.

[LGB+11] Mathias Lasgaard, Luc Goossens, Rikke Holm Bramsen, Tea Trillingsgaard, and Ask Elklit. Different sources of loneliness are associated with different forms of psychopathology in adolescence. *Journal of Research in Personality*, 45(2):233–237, 2011.

[LJF21] Leodoro J Labrague, Santos JAAD, and Charlie Falguera. Social and emotional loneliness among college students during the covid-19 pandemic: the predictive role of coping behaviours, social support, and personal resilience. 2021.

[LSWL16] Zhongqiu Li, Dianxi Shi, Feng Wang, and Fan Liu. Loneliness recognition based on mobile phone data. In *2016 International Symposium on Advances in Electrical, Electronics and Computer Engineering*. Atlantis Press, 2016.

[LTC14] Suyinn Lee, Cai Lian Tam, and Qiu Ting Chie. Mobile phone usage preferences: The contributing factors of personality, social anxiety and loneliness. *Social Indicators Research*, 118:1205–1228, 2014.

[LWGB15] Sanne MA Lamers, Gerben J Westerhof, Cees AW Glas, and Ernst T Bohlmeijer. The bidirectional relation between positive mental health and psychopathology in a longitudinal representative panel study. *The Journal of Positive Psychology*, 10(6):553–560, 2015.

[McA21] Kat J McAlpine. Depression, anxiety, loneliness are peaking in college students. *The Brink*, 17, 2021.

[MCHC11] Christopher M Masi, Hsi-Yuan Chen, Louise C Hawkley, and John T Cacioppo. A meta-analysis of interventions to reduce loneliness. *Personality and social psychology review*, 15(3):219–266, 2011.

[MCLP10] Anmol Madan, Manuel Cebrian, David Lazer, and Alex Pentland. Social sensing for epidemiological behavior change. In *Proceedings of the 12th ACM international conference on Ubiquitous computing*, pages 291–300, 2010.

[MCP12] David P MacKinnon, JeeWon Cheong, and Angela G Pirlott. Statistical mediation analysis. 2012.

[MJFS21] Laurie McLay, Hamish A Jamieson, Karyn G France, and Philip J Schluter. Loneliness and social isolation is associated with sleep problems among older community dwelling women and men with complex needs. *Scientific Reports*, 11(1):4877, 2021.

[ML16] Helen M Milojevich and Angela F Lukowski. Sleep and mental health in undergraduate students with generally healthy sleep habits. *PloS one*, 11(6):e0156372, 2016.

[MLC+20] Cathrine Mihalopoulos, Long Khanh-Dao Le, Mary Lou Chatterton, Jessica Bucholc, Julianne Holt-Lunstad, Michelle H Lim, and Lidia Engel. The economic costs of loneliness: a review of cost-of-illness and economic evaluation studies. *Social psychiatry and psychiatric epidemiology*, 55:823–836, 2020.

[MM99] Janet Morahan-Martin. The relationship between loneliness and internet use and abuse. *CyberPsychology & Behavior*, 2(5):431–439, 1999.

[MMS03] Janet Morahan-Martin and Phyllis Schumacher. Loneliness and social uses of the internet. *Computers in human behavior*, 19(6):659–671, 2003.

[MND+19] Marlies Maes, Stefanie A Nelemans, Sofie Danneel, Belén Fernández-Castilla, Wim Van den Noortgate, Luc Goossens, and Janne Vanhalst. Loneliness and social anxiety across childhood and adolescence: Multilevel meta-analyses of cross-sectional and longitudinal associations. *Developmental psychology*, 55(7):1548, 2019.

[MOEG17] Alicia Martinez, Virginia Ortiz, Hugo Estrada, and Miguel Gonzalez. A predictive model for automatic detection of social isolation in older adults. In *2017 International Conference on Intelligent Environments (IE)*, pages 68–75. IEEE, 2017.

[Mor13] William P Morgan. *Physical activity and mental health*. Taylor & Francis, 2013.

[MPP05] Nicholas Mays, Catherine Pope, and Jennie Popay. Systematically reviewing qualitative and quantitative evidence to inform management and policy-making in the health field. *Journal of health services research & policy*, 10(1_suppl):6–20, 2005.

[MPTK+20] Joanna McHugh Power, Jianjun Tang, Rose Ann Kenny, Brian A Lawlor, and Frank Kee. Mediating the relationship between loneliness and cognitive function: the role of depressive and anxiety symptoms. *Aging & mental health*, 24(7):1071–1078, 2020.

[MQLM22] Marlies Maes, Pamela Qualter, Gerine MA Lodder, and Marcus Mund. How (not) to measure loneliness: a review of the eight most commonly used scales. *International journal of environmental research and public health*, 19(17):10816, 2022.

[MRS+22] Tahsin Mullick, Ana Radovic, Sam Shaaban, Afsaneh Doryab, et al. Predicting depression in adolescents using mobile and wearable sensors: multimodal machine learning–based exploratory study. *JMIR Formative Research*, 6(6):e35807, 2022.

[MRTALPCS21] Alexandra Martín-Rodríguez, Jose Francisco Tornero-Aguilera, Pedro Javier López-Pérez, and Vicente Javier Clemente-Suárez. The effect of loneliness in psychological and behavioral profile among high school students in spain. *Sustainability*, 14(1):168, 2021.

[MSR+20] Muntazir Mehdi, Michael Stach, Constanze Riha, Patrick Neff, Albi Dode, Rüdiger Pryss, Winfried Schlee, Manfred Reichert, Franz J Hauck, et al. Smartphone and mobile health apps for tinnitus: systematic identification, analysis, and assessment. *JMIR mHealth and uHealth*, 8(8):e21767, 2020.

[MTOA+21] Isaac Moshe, Yannik Terhorst, Kennedy Opoku Asare, Lasse Bosse Sander, Denzil Ferreira, Harald Baumeister, David C Mohr, and Laura Pulkki-Råback. Predicting symptoms of depression and anxiety using smartphone and wearable data. *Frontiers in psychiatry*, 12:625247, 2021.

[MVVdNG17] Marlies Maes, Janne Vanhalst, Wim Van den Noortgate, and Luc Goossens. Intimate and relational loneliness in adolescence. *Journal of Child and Family Studies*, 26:2059–2069, 2017.

[NA18] Benjamin W Nelson and Nicholas B Allen. Extending the passive-sensing toolbox: using smart-home technology in psychological science. *Perspectives on psychological science*, 13(6):718–733, 2018.

[NGT+20] Nematjon Narziev, Hwarang Goh, Kobiljon Toshnazarov, Seung Ah Lee, Kyong-Mee Chung, and Youngtae Noh. Stdd: Short-term depression detection with passive sensing. *Sensors*, 20(5):1396, 2020.

[NNC18] Rebecca Nowland, Elizabeth A Necka, and John T Cacioppo. Loneliness and social internet use: pathways to reconnection in a digital world? *Perspectives on psychological science*, 13(1):70–87, 2018.

[OKDE+22] Blessing Ugochi Ojembe, Michael Ebe Kalu, Chigozie Donatus Ezulike, Anthony Obinna Iwuagwu, Prince Chiagozie Ekoh, Oluwagbemiga Oyinlola, Temitope Osifeso, John Osuolale Makanjuola, and Lydia Kapiriri. Understanding social and emotional loneliness among black older adults: A scoping review. *Journal of Applied Gerontology*, 41(12):2594–2608, 2022.

[oSoBS20] National Academies of Sciences, Division of Behavioral, and Social Sciences. *Social isolation and loneliness in older adults: Opportunities for the health care system.* National Academies Press, 2020.

[ÖTFM18] Fatma Özkan Tuncay, Tülay Fertelli, and Mukadder Mollaoğlu. Effects of loneliness on illness perception in persons with a chronic disease. *Journal of clinical nursing*, 27(7-8):e1494–e1500, 2018.

[OWZS10] Kelly O'Brien, Annette Wilkins, Elisse Zack, and Patricia Solomon. Scoping the field: identifying key research priorities in hiv and rehabilitation. *AIDS and Behavior*, 14(2):448–458, 2010.

[PA16] Gauri Pulekar and Emmanuel Agu. Autonomously sensing loneliness and its interactions with personality traits using smartphones. In *2016 IEEE Healthcare Innovation Point-Of-Care*

Technologies Conference (HI-POCT), pages 134–137. IEEE, 2016.

[PAK+13] Johanna Petersen, Daniel Austin, Jeffrey A Kaye, Misha Pavel, and Tamara L Hayes. Unobtrusive in-home detection of time spent out-of-home with applications to loneliness and physical activity. *IEEE journal of biomedical and health informatics*, 18(5):1590–1596, 2013.

[Pay08] Eugene S Paykel. Basic concepts of depression. *Dialogues in clinical neuroscience*, 10(3):279–289, 2008.

[PDA05] Marco Aurélio Monteiro Peluso and Laura Helena Silveira Guerra De Andrade. Physical activity and mental health: the association between exercise and mood. *Clinics*, 60(1):61–70, 2005.

[PIDWD20] Angela K Perone, Berit Ingersoll-Dayton, and Keisha Watkins-Dukhie. Social isolation loneliness among lgbt older adults: Lessons learned from a pilot friendly caller program. *Clinical Social Work Journal*, 48(1):126–139, 2020.

[PMG+20] Caroline Park, Amna Majeed, Hartej Gill, Jocelyn Tamura, Roger C Ho, Rodrigo B Mansur, Flora Nasri, Yena Lee, Joshua D Rosenblat, Elizabeth Wong, et al. The effect of loneliness on distinct health outcomes: a comprehensive review and meta-analysis. *Psychiatry Research*, 294:113514, 2020.

[PMHJ+21] Eiluned Pearce, Pamela Myles-Hooton, Sonia Johnson, Emily Hards, Samantha Olsen, Denisa Clisu, Sarah MA Pais, Heather A Chesters, Shyamal Shah, Georgia Jerwood, et al. Loneliness as an active ingredient in preventing or alleviating youth anxiety and depression: a critical interpretative synthesis incorporating principles from rapid realist reviews. *Translational psychiatry*, 11(1):628, 2021.

[PP81] Daniel Perlman and L Anne Peplau. Toward a social psychology of loneliness. *Personal relationships*, 3:31–56, 1981.

[PSS+17] Brian A Primack, Ariel Shensa, Jaime E Sidani, Erin O Whaite, Liu yi Lin, Daniel Rosen, Jason B Colditz, Ana Radovic, and Elizabeth Miller. Social media use and perceived social isolation among young adults in the us. *American journal of preventive medicine*, 53(1):1–8, 2017.

[PTAK16] Johanna Petersen, Stephen Thielke, Daniel Austin, and Jeffrey Kaye. Phone behaviour and its relationship to loneliness in older adults. *Aging & mental health*, 20(10):1084–1091, 2016.

[PvHH+19] Anubhuti Poudyal, Alastair van Heerden, Ashley Hagaman, Sujen Man Maharjan, Prabin Byanjankar, Prasansa Subba, and Brandon A Kohrt. Wearable digital sensors to identify risks of postpartum depression and personalize psychological treatment for adolescent mothers: protocol for a mixed methods exploratory study in rural nepal. *JMIR research protocols*, 8(9):e14734, 2019.

[QBR+13] Pamela Qualter, Stephen L Brown, Ken J Rotenberg, Janne Vanhalst, Rebecca A Harris, Luc Goossens, M Bangee, and P Munn. Trajectories of loneliness during childhood and adolescence: Predictors and health outcomes. *Journal of adolescence*, 36(6):1283–1293, 2013.

[QGY18] Juan Carlos Quiroz, Elena Geangu, and Min Hooi Yong. Emotion recognition using smart watch sensor data: Mixed-design study. *JMIR mental health*, 5(3):e10153, 2018.

[QM02] Pamela Qualter and Penny Munn. The separateness of social and emotional loneliness in childhood. *Journal of child psychology and psychiatry*, 43(2):233–244, 2002.

[QVH+15] Pamela Qualter, Janne Vanhalst, Rebecca Harris, Eeske Van Roekel, Gerine Lodder, Munirah Bangee, Marlies Maes, and Maaike Verhagen. Loneliness across the life span. *Perspectives on psychological science*, 10(2):250–264, 2015.

[QZWP23] Malik Muhammad Qirtas, Evi Zafeiridi, Eleanor Bantry White, and Dirk Pesch. The relationship between loneliness and depression among college students: Mining data derived from passive sensing. *Digital Health*, 9:20552076231211104, 2023.

[RCRY84] Dan Russell, Carolyn E Cutrona, Jayne Rose, and Karen Yurko. Social and emotional loneliness: an examination of weiss's typology of loneliness. *Journal of personality and social psychology*, 46(6):1313, 1984.

[Rei21] Harry T Reis. Gender effects in social participation: Intimacy, loneliness, and the conduct of social interaction. In *The emerging field of personal relationships*, pages 91–105. Routledge, 2021.

[RLHE18] Dmitri Rozgonjuk, Jason C Levine, Brian J Hall, and Jon D Elhai. The association between problematic smartphone use, depression and anxiety symptom severity, and objectively measured smartphone use over one week. *Computers in Human Behavior*, 87:10–17, 2018.

[RLKS06] Ami Rokach, Rachel Lechcier-Kimel, and Artem Safarov. Loneliness of people with physical disabilities. *Social Behavior and Personality: an international journal*, 34(6):681–700, 2006.

[RMP+23] Jessica Rees, Faith Matcham, Freya Probst, Sebastien Ourselin, Yu Shi, Michela Antonelli, Anthea Tinker, and Wei Liu. Wearables, sensors and the future of technology to detect and infer loneliness in older adults. *Gerontechnology*, 22(2), 2023.

[Rok12] Ami Rokach. Loneliness updated: An introduction. *The Journal of psychology*, 146(1-2):1–6, 2012.

[Rok19] Ami Rokach. *The psychological journey to and from loneliness: Development, causes, and effects of social and emotional isolation.* Academic Press, 2019.

[ROKvdM+12] Nathaly Rius-Ottenheim, Daan Kromhout, Roos C van der Mast, Frans G Zitman, Johanna M Geleijnse, and Erik J Giltay. Dispositional optimism and loneliness in older men. *International journal of geriatric psychiatry*, 27(2):151–159, 2012.

[Roy83] J Patrick Royston. Some techniques for assessing multivarate normality based on the shapiro-wilk w. *Journal of the Royal Statistical Society Series C: Applied Statistics*, 32(2):121–133, 1983.

[RPC80] Dan Russell, Letitia A Peplau, and Carolyn E Cutrona. The revised ucla loneliness scale: concurrent and discriminant validity evidence. *Journal of personality and social psychology*, 39(3):472, 1980.

[RPP+20] Alessandro Rossi, Anna Panzeri, Giada Pietrabissa, Gian Mauro Manzoni, Gianluca Castelnuovo, and Stefania Mannarini. The anxiety-buffer hypothesis in the time of covid-19: when self-esteem protects from the impact of loneliness and fear on anxiety and depression. *Frontiers in psychology*, 11:2177, 2020.

[RS+22] Eka Mala Sari Rochman, Herry Suprajitno, et al. Comparison of clustering in tuberculosis using fuzzy c-means and k-means methods. *Commun. Math. Biol. Neurosci.*, 2022:Article–ID, 2022.

[Rus96] Daniel W Russell. Ucla loneliness scale (version 3): Reliability, validity, and factor structure. *Journal of personality assessment*, 66(1):20–40, 1996.

[SANA20] Anna Sundström, Annelie Nordin Adolfsson, Maria Nordin, and Rolf Adolfsson. Loneliness increases the risk of all-cause dementia and alzheimer's disease. *The Journals of Gerontology: Series B*, 75(5):919–926, 2020.

[SANV+22] Fatemeh Sarhaddi, Iman Azimi, Hannakaisa Niela-Vilén, Anna Axelin, Pasi Liljeberg, and Amir M Rahmani. Predicting maternal social loneliness by passive sensing with wearable devices. *medRxiv*, pages 2022–10, 2022.

[SD14] Gillian M Sandstrom and Elizabeth W Dunn. Social interactions and well-being: The surprising power of weak ties. *Personality and Social Psychology Bulletin*, 40(7):910–922, 2014.

[SFJ+05] Joshua R Smith, Kenneth P Fishkin, Bing Jiang, Alexander Mamishev, Matthai Philipose, Adam D Rea, Sumit Roy, and Kishore Sundara-Rajan. Rfid-based techniques for human-activity detection. *Communications of the ACM*, 48(9):39–44, 2005.

[SH21] Amanda Sundqvist and Jessica Hemberg. Adolescents' and young adults' experiences of loneliness and their thoughts about its alleviation. *International Journal of Adolescence and Youth*, 26(1):238–255, 2021.

[SHC+21] Natasja Schutter, Tjalling J Holwerda, Hannie C Comijs, Paul Naarding, Rien HL Van, Jack JM Dekker, Max L Stek, and Didi Rhebergen. Loneliness, social network size, and mortality in older adults and the role of cortisol. *Aging & mental health*, 25(12):2246–2254, 2021.

[SHGDF20] Bridget Shovestul, Jiayin Han, Laura Germine, and David Dodell-Feder. Risk factors for loneliness: The high relative importance of age versus other factors. *PloS one*, 15(2):e0229087, 2020.

[SJD+23] Sarah Schroyen, N Janssen, LA Duffner, M Veenstra, Effrosyni Pyrovolaki, Eric Salmon, Stéphane Adam, et al. Prevalence of loneliness in older adults: a scoping review. *Health & Social Care in the Community*, 2023, 2023.

[SK13] Bhawana Singh and UV Kiran. Loneliness among elderly women. *International Journal of Humanities and Social Science Invention*, 2(2):10–14, 2013.

[SMC+15] Wendy Sanchez, Alicia Martinez, Wilfrido Campos, Hugo Estrada, and Vicente Pelechano. Inferring loneliness levels in older adults from smartphones. *Journal of Ambient Intelligence and Smart Environments*, 7(1):85–98, 2015.

[SNČS21] Miloš Stanković, Milkica Nešić, Svetlana Čičević, and Zhuanghua Shi. Association of smartphone use with depression, anxiety, stress, sleep quality, and internet addiction. empirical evidence from a smartphone application. *Personality and individual differences*, 168:110342, 2021.

[SP22] Christine Stephens and Hannah Phillips. Older people's neighborhood perceptions are related to social and emotional loneliness and mediated by social network type. *The Gerontologist*, 62(9):1336–1346, 2022.

[SQK21] Mahsa Sheikh, Meha Qassem, and Panicos A Kyriacou. Wearable, environmental, and smartphone-based passive sensing for mental health monitoring. *Frontiers in Digital Health*, 3:662811, 2021.

[SRRM+17] Sandra Servia-Rodríguez, Kiran K Rachuri, Cecilia Mascolo, Peter J Rentfrow, Neal Lathia, and Gillian M Sandstrom. Mobile sensing at the service of mental well-being: a large-scale longitudinal study. In *Proceedings of the 26th International Conference on World Wide Web*, pages 103–112, 2017.

[SVA+22] Aditi Site, Saigopal Vasudevan, Samuel Olaiya Afolaranmi, Jose L Martinez Lastra, Jari Nurmi, and Elena Simona Lohan. A machine-learning-based analysis of the relationships between loneliness metrics and mobility patterns for elderly. *Sensors*, 22(13):4946, 2022.

[SY20] Kristina P Sinaga and Miin-Shen Yang. Unsupervised k-means clustering algorithm. *IEEE access*, 8:80716–80727, 2020.

[SZK+15] Sohrab Saeb, Mi Zhang, Christopher J Karr, Stephen M Schueller, Marya E Corden, Konrad P Kording, David C Mohr, et al. Mobile phone sensor correlates of depressive symptom severity in daily-life behavior: an exploratory study. *Journal of medical Internet research*, 17(7):e4273, 2015.

[THH11] Sara Thomée, Annika Härenstam, and Mats Hagberg. Mobile phone use and stress, sleep disturbances, and symptoms of depression among young adults-a prospective cohort study. *BMC public health*, 11(1):1–11, 2011.

[TK14] Greg Townley and Bret Kloos. Mind over matter? the role of individual perceptions in understanding the social ecology of housing environments for individuals with psychiatric disabilities. *American journal of community psychology*, 54(3-4):205–218, 2014.

[TLZ+18] Andrea C Tricco, Erin Lillie, Wasifa Zarin, Kelly K O'Brien, Heather Colquhoun, Danielle Levac, David Moher, Micah DJ Peters, Tanya Horsley, Laura Weeks, et al. Prisma extension for scoping reviews (prisma-scr): checklist and explanation. *Annals of internal medicine*, 169(7):467–473, 2018.

[TMSA$^+$22] Thamara Tapia-Muñoz, Ursula M Staudinger, Kasim Allel, Andrew Steptoe, Claudia Miranda-Castillo, José T Medina, and Esteban Calvo. Income inequality and its relationship with loneliness prevalence: A cross-sectional study among older adults in the us and 16 european countries. *Plos one*, 17(12):e0274518, 2022.

[TOO$^+$19] Alina Trifan, Maryse Oliveira, José Luís Oliveira, et al. Passive sensing of health outcomes through smartphones: systematic review of current solutions and possible limitations. *JMIR mHealth and uHealth*, 7(8):e12649, 2019.

[TR21] Saurabh Singh Thakur and Ram Babu Roy. Predicting mental health using smart-phone usage and sensor data. *Journal of Ambient Intelligence and Humanized Computing*, 12(10):9145–9161, 2021.

[UKT$^+$18] Md Uddin, Weria Khaksar, Jim Torresen, et al. Ambient sensors for elderly care and independent living: a survey. *Sensors*, 18(7):2027, 2018.

[UUHL99] Bert N Uchino, Darcy Uno, and Julianne Holt-Lunstad. Social support, physiological processes, and health. *Current Directions in Psychological Science*, 8(5):145–148, 1999.

[VAAB$^+$16] C VandeWeerd, Ngozichukwuka Agu, Garrick Aden-Buie, Chad Radwan, Kimberly Lent, Efe Yetisener, E Zapf, C Watkins, M Roberts, and A Yalcin. Ambient sensor monitoring: Detecting loneliness and depression in older adults via passive home sensing. In *APHA 2016 Annual Meeting & Expo (Oct. 29-Nov. 2, 2016)*. American Public Health Association, 2016.

[VAIF$^+$22] Barbara Adriana Lambert Van As, Enrico Imbimbo, Angela Franceschi, Ersilia Menesini, and Annalaura Nocentini. The longitudinal association between loneliness and depressive symptoms in the elderly: a systematic review. *International Psychogeriatrics*, 34(7):657–669, 2022.

[Vau88] Alan Vaux. Social and emotional loneliness: The role of social and personal characteristics. *Personality and Social Psychology Bulletin*, 14(4):722–734, 1988.

[VdLPB03] Mark Van der Laan, Katherine Pollard, and Jennifer Bryan. A new partitioning around medoids algorithm. *Journal of Statistical Computation and Simulation*, 73(8):575–584, 2003.

[VHTC11] Tyler J VanderWeele, Louise C Hawkley, Ronald A Thisted, and John T Cacioppo. A marginal structural model analysis for loneliness: implications for intervention trials and clinical practice. *Journal of consulting and clinical psychology*, 79(2):225, 2011.

[Vic23] Shunya Vichaar. Incremental learning in xgboost. `https://shunya-vichaar.medium.com/incremental-learning-in-xgboost-b3eac6135ce`, 2023.

[VLA+20] J Vega, M Li, K Aguillera, N Goel, E Joshi, KC Durica, AR Kunta, and CA Low. Rapids: Reproducible analysis pipeline for data streams collected with mobile devices. *J Med Internet Res Preprints. URL: https://preprints. jmir. org/preprint/23246 [accessed 2020-08-18]*, 2020.

[VLS+13] Janne Vanhalst, Koen Luyckx, Ron HJ Scholte, Rutger CME Engels, and Luc Goossens. Low self-esteem as a risk factor for loneliness in adolescence: Perceived-but not actual-social acceptance as an underlying mechanism. *Journal of abnormal child psychology*, 41:1067–1081, 2013.

[VP07] Patti M Valkenburg and Jochen Peter. Online communication and adolescent well-being: Testing the stimulation versus the displacement hypothesis. *Journal of Computer-Mediated Communication*, 12(4):1169–1182, 2007.

[WBC+21] Congyu Wu, Amanda N Barczyk, R Cameron Craddock, Gabriella M Harari, Edison Thomaz, Jason D Shumake, Christopher G Beevers, Samuel D Gosling, and David M Schnyer. Improving prediction of real-time loneliness and companionship type using geosocial features of personal smartphone data. *Smart Health*, 20:100180, 2021.

[WBK+20] Juliet RH Wakefield, Mhairi Bowe, Blerina Kellezi, Andrew Butcher, and John A Groeger. Longitudinal associations between family identification, loneliness, depression, and sleep quality. *British Journal of Health Psychology*, 25(1):1–16, 2020.

[WC23] Katarzyna Wolska and Ann-Marie Creaven. Associations between transient and chronic loneliness, and depression, in the understanding society study. *British Journal of Clinical Psychology*, 62(1):112–128, 2023.

[WCC+14] Rui Wang, Fanglin Chen, Zhenyu Chen, Tianxing Li, Gabriella Harari, Stefanie Tignor, Xia Zhou, Dror Ben-Zeev, and Andrew T Campbell. Studentlife: assessing mental health, academic performance and behavioral trends of college students using smartphones. In *Proceedings of the 2014 ACM international joint conference on pervasive and ubiquitous computing*, pages 3–14, 2014.

[WCT88] David Watson, Lee Anna Clark, and Auke Tellegen. Development and validation of brief measures of positive and negative affect: the panas scales. *Journal of personality and social psychology*, 54(6):1063, 1988.

[Wei75] Robert Weiss. *Loneliness: The experience of emotional and social isolation*. MIT press, 1975.

[WKL+14] Lorcan Walsh, Andrea Kealy, John Loane, Julie Doyle, and Rodd Bond. Inferring health metrics from ambient smart home data. In *2014 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, pages 27–32. IEEE, 2014.

[WN12] John Watson and Drew Nesdale. Rejection sensitivity, social withdrawal, and loneliness in young adults. *Journal of Applied Social Psychology*, 42(8):1984–2005, 2012.

[WPA+22] Nutchar Wiwatkunupakarn, Chanapat Pateekhum, Chanchanok Aramrat, Wichuda Jiraporncharoen, Kanokporn Pinyopornpanish, and Chaisiri Angkurawaranon. Social networking site usage: A systematic review of its relationship with social isolation, loneliness, and depression among older adults. *Aging & mental health*, 26(7):1318–1326, 2022.

[WPB+21] Britta Wetzel, Rüdiger Pryss, Harald Baumeister, Johanna-Sophie Edler, Ana Sofia Oliveira Gonçalves, and Caroline

Cohrdes. "how come you don't call me?" smartphone communication app usage as an indicator of loneliness and socialwell-being across the adult lifespan during the covid-19 pandemic. *International Journal of Environmental Research and Public Health*, 18(12):6212, 2021.

[WS16] Qiang Wu and Dave Snell. An introduction to incremental learning. *Predictive Analytics and Futurism*, 13:18–19, 2016.

[WSS+13] Fang Wang, Erik Stone, Marjorie Skubic, James M Keller, Carmen Abbott, and Marilyn Rantz. Toward a passive low-cost in-home gait assessment system for older adults. *IEEE journal of Biomedical and Health Informatics*, 17(2):346–355, 2013.

[XB91] Xuanli Lisa Xie and Gerardo Beni. A validity measure for fuzzy clustering. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, 13(08):841–847, 1991.

[XNG14] Qianli Xu, Tin Lay Nwe, and Cuntai Guan. Cluster-based analysis for personalized stress evaluation using physiological signals. *IEEE journal of biomedical and health informatics*, 19(1):275–281, 2014.

[XT15] Dongkuan Xu and Yingjie Tian. A comprehensive survey of clustering algorithms. *Annals of data science*, 2:165–193, 2015.

[XZS+22] Xuhai Xu, Han Zhang, Yasaman Sefidgar, Yiyi Ren, Xin Liu, Woosuk Seo, Jennifer Brown, Kevin Kuehn, Mike Merrill, Paula Nurius, et al. Globem dataset: Multi-year datasets for longitudinal human behavior modeling generalization. *Advances in Neural Information Processing Systems*, 35:24655–24692, 2022.

[YAJ+23] Zhongqi Yang, Iman Azimi, Salar Jafarlou, Sina Labbaf, Jessica Borelli, Nikil Dutt, and Amir M Rahmani. Loneliness forecasting using multi-modal wearable and mobile sensing in everyday settings. In *2023 IEEE 19th International Conference on Body Sensor Networks (BSN)*, pages 1–4. IEEE, 2023.

[YFLL20] Jiaxin Yang, Xi Fu, Xiaoli Liao, and Yamin Li. Association of problematic smartphone use with poor sleep quality, depression, and anxiety: A systematic review and meta-analysis. *Psychiatry research*, 284:112686, 2020.

[YLH$^+$14] W Quin Yow, Xiaoqian Li, Wan-Yu Hung, Megan Goldring, Long Cheng, and Yu Gu. Predicting social networks and psychological outcomes through mobile phone sensing. In *2014 IEEE international conference on communications (ICC)*, pages 3925–3931. IEEE, 2014.

[Zam08] Vanda Lucia Zammuner. Italians-social and emotional loneliness: the results of five studies. *International Journal of Educational and Pedagogical Sciences*, 2(4):416–428, 2008.

[ZJ21] Samuele Zilioli and Yanping Jiang. Endocrine and immunomodulatory effects of social isolation and loneliness across adulthood. *Psychoneuroendocrinology*, 128:105194, 2021.