Objective Monitoring of Loneliness Levels using Smart Devices: A Multi-Device Approach for Mental Health Applications

Salar Jafarlou[a], Iman Azimi[a,b], Jocelyn Lai[c], Yuning Wang[d], Sina Labbaf[a], Brenda Nguyen[c], Hana Qureshi[c], Christopher Marcotullio[c], Jessica L. Borelli[e], Nikil D. Dutt[a], Amir M. Rahmani[a,b,f]

[a]*Donald Bren School of Information and Computer Sciences, University of California, Irvine*
[b]*Institute for Future Health, University of California, Irvine*
[c]*Department of Psychological Science, University of California, Irvine*
[d] *Department of Computing, University of Turku, Turku, Finland.*
[e]*Department of Cognitive Science, University of California, Irvine*
[f]*School of Nursing, University of California, Irvine*

Abstract

Loneliness is linked to wide ranging physical and mental health problems, including increased rates of mortality. Understanding how loneliness manifests is important for targeted public health treatment and intervention. With advances in mobile sending and wearable technologies, it is possible to collect data on human phenomena in a continuous and uninterrupted way. In doing so, such approaches can be used to monitor physiological and behavioral aspects relevant to an individual's loneliness. In this study, we proposed a method for continuous detection of loneliness using fully objective data from smart devices and passive mobile sensing. We also investigated whether physiological and behavioral features differed in their importance in predicting loneliness across individuals. Finally, we examined how informative data from each device is for loneliness detection tasks. We assessed subjective feelings of loneliness while monitoring behavioral and physiological patterns in 30 college students over a 2-month period. We used smartphones to monitor behavioral patterns (e.g., location changes, type of notifications, in-coming and out-going calls/text messages) and smart watches and rings to monitor physiology and sleep patterns (e.g., heart-rate, heart-rate variability, sleep duration). We also collected participants' loneliness feeling scales multiple times a day through a questionnaire app on their phone. Using the data collected from their devices, we trained a random forest machine learning based model to detect loneliness levels. We found support for loneliness prediction using a multi-device and fully-objective approach. Furthermore, behavioral data collected by smartphones generally were the most important features across all participants. The study provides promising results for using objective data to monitor mental health indicators, which could provide a continuous and uninterrupted source of information in mental healthcare applications.

Keywords: wearable devices; mental health; affective computing

1. Introduction

1.1. Loneliness and Social Isolation – a Public Health Epidemic

Though frequently used and described in the same contexts, loneliness is distinct from social isolation. Loneliness refers to a subjective feeling wherein the individual experiences a lack of actual connection within social interactions [1, 2]. Social isolation refers to the specific lack of social interaction or access to one's social networks [3]. Both loneliness and social isolation have profound health implications and are linked with increased morbidity and mortality rates [4, 5]. Concerns for the health implications of loneliness and social isolation were heightened during the COVID-19 pandemic [6, 7], during which people around the world were requested to minimize contact and comply with stay-at-home orders to reduce transmission of the disease. Individuals with lower socioeconomic status, younger adults, and students in particular were found to have high feelings of loneliness during the pandemic [8, 9], though empirical evidence regarding the

differences in loneliness over time and the psychological implications of lockdown remain unclear [10, 11, 12]. However, a recent meta-analysis by Ernst and colleagues [13] suggested a small yet significant effect size in the differences between feelings of loneliness before and after the pandemic. Several researchers argue that there are differential associations between loneliness and social isolation with health outcomes, such that loneliness contributes more to mental health outcomes while social isolation is more important in predicting declines in cognitive and physical health [14, 15, 16, 17]. However, other studies suggest that the differential health implications of loneliness and social isolation are much more difficult to disentangle; for example both constructs are associated with increased risk for cardiovascular disease and interact in their relation with mortality [18, 3, 19]. In summary, given their crucial link to health-risks, both loneliness and social isolation are necessary to address. To target ways of reducing these feelings, it is important to first determine when and how they occur.

1.2. Assessments and Predictions of Loneliness and Social Isolation in Daily Life

Although a robust body of literature has established a strong link between loneliness and health outcomes [20, 21], this work is constrained by the fact that the majority of this work has relied on self-reported measures of loneliness and social isolation that require that participants report on their symptoms spanning an extended period of time (e.g., in the past two weeks). It can be difficult to accurately report on symptoms over such a long time period, and such symptom reports can be influenced by current mood [22, 23, 24]. An on-going and exciting avenue to examine loneliness, both its occurrence and association with health and social outcomes, is through its occurrence in daily life [25]. By assessing loneliness with greater ecological validity, we may be able to understand how and when loneliness plays a role in physical and mental health outcomes. For example, existing literature has found that loneliness in daily life is linked to increased feelings of negative affect, greater negative appraisals of social interactions, and more time spent alone; furthermore, these associations had lagged associations with loneliness and depression at a later time point [26, 25].

In addition to outcomes associated with loneliness, several studies have been conducted to assess the feasibility of detecting and modeling loneliness and social isolation in real-world settings. Previous approaches using intensive longitudinal self-reports can put a heavier load on users, resulting in unsatisfactory experiences for the users due to the reliance on frequent engagement [27, 28]. An attempt to resolve this issue involves leveraging users' behavioral data as collected by continuous mobile sensing. Using continuous and objective sensing devices (e.g., smartwatches) to monitor loneliness allows for a more accurate and representative assessment when taking into account contextual factors in a user's environment. Smartphones can monitor certain aspects of a user's life using multiple embedded sensors (e.g. GPS, accelerometer, microphone, light sensor and phone usage logs). Furthermore, passive sensing can quantify the behavioral patterns of participants, including the type and frequency of activities performed, changes in geographic location, and phone usage. Such work has already begun to leverage this type of data to determine features that contribute to loneliness with high predictive accuracy [29, 30, 31, 32]. For example, mobility and digital media use are found to be features linked to feelings of loneliness [33, 27]. Additionally, the presence of other people and differences in phone usage were also shown to be predictors of loneliness [33, 34]. Doryab and colleagues [33] have attempted to classify loneliness levels amongst college students before and after a semester by leveraging data collected with smart wristbands and phones. Using these devices, Doryab and colleagues were able to extract features about calls, locations and screen usage collected from phones along with sleep and activity level features from smart wristbands. Nevertheless, the loneliness scale labels obtained in this study were limited to only two timepoints: pre- and post-semester, which obscures the temporal dynamics of the loneliness data. The wristband used in this study (Fitbit Flex 2) assessed sleep information using the same motion sensors used for activity level recognition. This is while recent devices are capable of leveraging heart rate and other physiological sensors for sleep quality assessment [35, 36].

Despite existing literature that has leveraged the capabilities of smartphones in monitoring loneliness and well-being more broadly, a few novel directions in this approach have yet to be fully explored. First, ubiquitous sensing and advances in technology can allow us to potentially monitor and predict loneliness with higher granularity. Most studies have assessed loneliness as a trait measure or with few assessments throughout the study duration (Bello-Valle et al., 2022; Doryab et al., 2018; Gao et al., 2016; MacDonald

et al., 2020) as opposed to intensive assessments of loneliness across multiple days over a long period of time. For example, work by [34] and [37] measured loneliness, social anxiety, and affect at the beginning and end of their study. Although measuring these levels at two time points can provide general trends, such approaches likely omit the temporal changes in loneliness during the course of the study or at the daily level. There are, however, a few studies that have assessed loneliness and mental health states multiple times a day [38, 39]. In a study by Ben-Zeev et. al. [38], participants completed a 10-week study that passively collected their speech patterns, location, kinesthetic activities, and sleep duration using smartphones alongside self-reported assessments of daily stress, depression, and loneliness. StudentLife [39], was another study conducted by Wang and colleagues [39], participating 48 students during a 10 week period collecting various contextual data from phone and comprehensive ecological momentary assessments (EMA) assessing variables including stress, mood, and loneliness etc. The authors performed detailed analysis on the correlation of contextual information extracted by smartphones and different EMA data however they did not report loneliness detection/classification tasks. While these works provide a relational analysis between the loneliness scales and data, our study aims to take a predictive approach to loneliness.

Lastly, an exciting avenue for the field is the use of a comprehensive approach to monitor and predict loneliness and mental states that involves multiple modalities of passive sensing. A limitation of previous work that relies on passive sensing using smartphones is the expectation of continuous user engagement. While individuals do frequently access their smartphones, there may be times where data collection through smartphones may lack granularity in assessment (e.g., leaving smartphones in bags while at the gym or doing an activity). With the progression of technological tools for entertainment, users, particularly younger adults, may use other modalities for their online presence, such as tablets, game consoles, and smartwatches. Furthermore, smartphone assessments are often limited to assessment of behaviors, unlike other ubiquitous devices such as smart wearables that can provide physiological state measurements. Indeed, prior literature that has used smartphones in real time typically focuses on usage of phone features (e.g., incoming and outgoing calls or messages), geolocation, and participant self-reported responses to assess how these behaviors relate to health and social outcomes [37, 39]. Incorporating other ubiquitous devices such as the OuraRing and Samsung smartwatch allows for the continuous collection of several features such as sleep (e.g., sleep duration, quality, restlessness) and cardiovascular (e.g., heart-rate, heart-rate variability, blood pressure) indicators [35]. The ability to assess sleep and cardiovascular indicators is important, given existing literature suggesting the links between loneliness with sleep disturbances and lower cardiac output [40, 26, 41]. Therefore, leveraging passive sensing using multiple devices with the capacity to detect unique factors such as sleep and cardiovascular activity would build upon existing literature to provide a more comprehensive and granulary understanding of when and how loneliness occurs.

1.3. Current Research

Altogether, the intersection of mental health and technology offers an exciting opportunity to build upon foundational work in the field [33, 39, 42] by developing a more comprehensive monitoring framework capable of assessing the occurrence and predictability of loneliness. In this study, we develop a method of passively and continuously monitoring loneliness using smart wearables and smartphone devices. Using these devices together allows us to assess a wide range of behavioral patterns such as phone usage, communications, and locations as well as physiological patterns such as sleep indices and cardiovascular data. We leverage these devices to propose and evaluate a fully objective loneliness detection method that does not rely on user engagement. We also aim to examine which behavioral and physiological features were most predictive of loneliness for each individual. Finally, our third aim is to determine device efficiency of loneliness monitoring and detection.

2. Method

2.1. Participants

Full-time college students (N=30) at a West coast university between the ages of 18-22 were recruited to participate in an intensive longitudinal study investigating loneliness and mental health . Participants were

eligible if they spoke fluent English and used an Android smartphone compatible with the Oura Ring and Samsung Active 2 watch. Students were ineligible to participate if they were married, had children, were returning to school after a three year hiatus, or if they experienced any severe forms of psychopathology (i.e., diagnosed with clinical depression or substance use disorders, psychosis, or any form of suicidal ideation). The exclusion criteria were intended to ensure a relatively homogenous sample of college students and individuals in emerging adulthood. We reasoned that students whose demographics were different (parents, older adults, and individuals returning to school) may have a different experience than emerging adults as they develop their identity and social networks . Students were recruited through faculty outreach where professors may then make announcements to their classes. Social media posts on institutional platforms (e.g., Facebook, Reddit, and Discord) were also used to recruit students. Interested participants reached out via email and were administered a screening survey to assess for moderate to severe depression or suicidal ideation. Individuals who met these criteria were then reached out by the clinical psychologist on our team (JB) for additional support and follow-up. Campus and local wellness resources were provided to all individuals who reached out or completed the screening survey.

2.2. Procedure

All procedures were approved by the researchers' academic institutional review board (#2019-5153). For the purpose of this investigation, only relevant procedures and our monitoring phase will be described from the larger study. If participants met the inclusion criteria, they were then contacted to schedule their first in-person lab session in which they completed a baseline battery assessment of mental health, emotions, and well-being related measures. The mental health assessment included measures of psychopathology, on the basis of which participants were withdrawn from the study if the research team determined in consultation with JB that continued participation would be unsafe. During the first lab session, Oura Ring, Samsung Gear Sport watch, and smartphone apps related to the study were set up and downloaded for the participants. Participants were then instructed on how to use the devices and were guided through tasks that they would be completing for the next four weeks of their participation. Participants were instructed to wear their devices at all times throughout the study period, except under certain conditions (e.g., while charging devices or while performing intensive activities that risked damage to the watch). For the smartphone apps, participants downloaded AWARE that monitored their phone usage and a study-designed app, mSavorUs, which would prompt individuals to complete a brief survey five times per day, using an interval-based ecological momentary assessment (EMA) design. The survey included questions about affect, interactions with others, and feelings of loneliness, social isolation, and connectedness. Apps relating to the wearable devices were also downloaded onto participants' phones for proper device functionality. Participants completed this monitoring phase for about 8-weeks and were compensated for their involvement in all components of the study.

2.3. Data Collection

To support our longitudinal study design, we built a platform capable of collecting and storing data efficiently. We designed a research dashboard that allowed the researchers to have constant access to the data being collected in order to monitor participants' progress. The dashboard visually displayed the collected data and a summary of each participant's daily activities. The information was then used to track and assess possible connectivity issues with the devices throughout the study duration. The collected data fell under three categories: objective physiological data, objective behavioral data, and subjective self-report questionnaires. Participants' objective physiological data was collected using the Oura ring and Samsung watch. The Oura ring was used to assess information regarding the participant's sleep patterns, and the Samsung watch was employed to conduct continuous physiological assessments throughout the day. The objective behavioral data was collected using the AWARE phone application. AWARE provides us with mobile sensory data and participants' phone usage details, which can then be used to derive behavioral features. Subjective self-report data was collected using the mSavorUs phone application that was developed for the purpose of collecting subjective experiential data and delivering interventions in this study.

2.3.1. Objective Physiological Data using Samsung Watch and Oura Ring

The Oura Ring offers a range of metrics regarding the user's physical activity, and sleep patterns [7]. However, Photoplethysmography (PPG)-based wearable devices, including the Oura Ring, can encounter noise during signal collection, particularly when utilized in a home-based monitoring setting. For our purposes, we used data from Oura that included the user's sleep, physiology and physical activity. The final list of selected data included heart-rate related information during sleep, the estimated time of start of sleep after getting into bed, as well as the number of minutes with high, medium and low level of physical activity during the day (refer to Appendix Table A.3 for a detailed list of features of Oura ring used in this study). Samsung Active 2 watch operated on the Tizen open-source operating system [43], allowing for the development of custom data collection applications. This open-source platform enabled us to develop an application specifically designed for the watch to collect the data in a manner customized to the purposes of our study. The watch app activated every two hours to collect twelve minutes of PPG signals, used to extract heart rate and heart rate variability. The collected data was then synchronized with our server through wifi or bluetooth.

2.3.2. Objective Behavioral Data using the AWARE app and Subjective Experiential Data using mSavorUs

For the objective behavioral data, we used the AWARE app available for Android. This app's sensors were configured to collect participants screen usage (e.g., lock, unlock time), application usage (e.g., notification, time of opening any app), their keyboard strokes pattern, battery level, communications (e.g., message/calls sending and receiving time) and GPS location [44]. Finally, for subjective experience, we designed and developed the mSavorUs application as a portion of a larger intervention-based study. This application [45] included assessments of loneliness and social isolation using mobile app interfaces and push notifications.

2.4. Data Analytic Plan

2.4.1. Feature Extraction

We developed Heart Rate (HR) and Heart Rate Variability (HRV) extraction methods from the raw PPG signals collected by the Samsung watch. The raw PPG signals are highly susceptible to environmental noise motion artifacts. To illustrate the problem, we plot two 60-second PPG signals. Figure 1 (a) is a clean PPG signal, showing heart beat oscillations. However, Figure 1(b) is a noisy PPG, distorted by the subject's hand movements. Such distorted signals will result in unreliable HR and HRV features extraction. Therefore, we developed a PPG processing pipeline to address this problem.

The pipeline includes 3 main stages: signal quality assessment (SQA), signal reconstruction, and PPG peak detection. The SQA classifies PPG signals as "clean" or "noisy" by extracting five features from the signal, including interquartile range, standard deviation of the power spectral density, range of energy of heart cycles, average Euclidean distances, and average correlation between a template and heart cycles [46]. After we performed SQA, short-term "noisy" segments (less than 15 seconds) were reconstructed using a trained generative adversarial network (GAN) model [47]. The GAN model was trained to reconstruct noisy PPG using the information both in the distorted part and its proceeding clean signals. Then, a trained Dilated Convolution Neural Network (DCNN) was employed to detect the systolic peaks [48] and inter-beat intervals (IBI). Finally, HR and HRV-related features were extracted from the IBI signals. We access the Oura ring data through an application programming interface (API) that provides the processed features of sleep and activity. No further feature extraction or post-processing steps were done on the ring data (please refer to Appendix Tables A.3, A.4 for the full list and description of the features extracted from ring and watch).

We developed methods to extract behavioral parameters of the participants' daily living through the longitudinal behavioral data collected by the smartphones. For calls, we used the duration sums and number of calls in each category (e.g., outgoing, incoming and voicemail). For location data we first located the participant's house as the most frequently visited place during the nights of study. Then given a location time window, we extracted variance of latitude, variance of speed, mean of speed, number of places, home duration, outdoor duration, mean of outdoor duration (>=2 places), standard deviation of outdoor duration

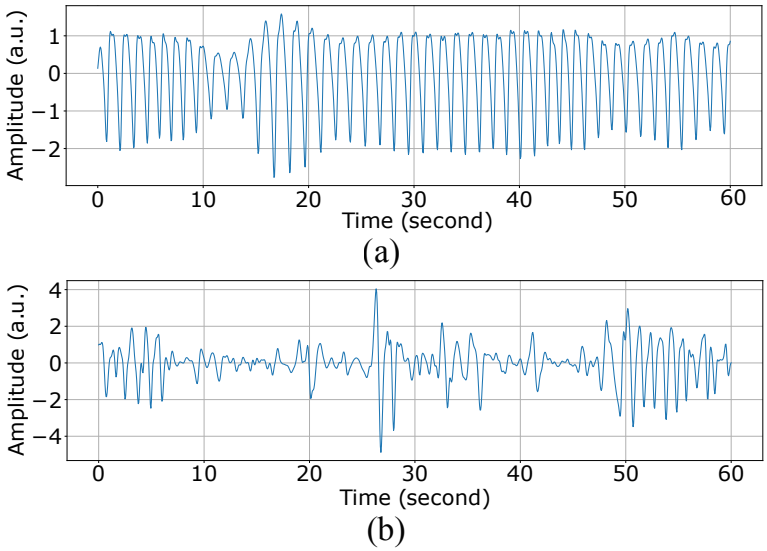

Figure 1: 60-second PPG segments (a) clean (b) distorted by hand movements

(>= 2 places), longest duration type other than home (>= 2 places), and total travel distance(>= 2 places). For messages and notifications, we counted the number of messages in each category. Table 1 demonstrates all of the categories and three samples for each of them (please refer to the Appendix Table A.3 for a detailed list of all features extracted from the phone.)

Table 1: Different categories of App notifications and with three examples of each

Apps Category	**App Examples**		
Productivity	Google Doc	Office Outlook	Gmail
Photography	Android Photos	Bazaar	Lightroom
Communication	Whatsapp	Android Messenger	Discord
Lifestyle	Oneconnect	T-Mobile Tuesdays	Samsung pay
Auto & Vehicles	Gearhead	Toyota	Carfax
Travel & Local	Couch Surfing	Google Maps	Uber
Education	duolingo	Wonder	Canvas
Finance	mint	American Express	SplitwiseMobile
Video Players & Editors	Vlc player	Youtube	Capcut
Social	Snapchat	Katana	TikTok
Books & Reference	audible	Scribd	Chirp
Shopping	Slickdeals	Amazon	Target
Health & Fitness	Fitindex	Myfitnesspal	shealth
Entertainment	Amazon Prime Video	Netflix	Hulu
Business	Duo	Slack	LinkedIn
Music & Audio	Spotify	Youtube Music	Soundcloud
Tools	Sprint	Android search box	Google Mobile Service
Unknown	Gametools	Clock	Calendar

Some of the features from the devices were collected in a single assessment per day (i.e., sleep related features) whereas others were assessed with higher resolution, multiple times throughout the day (i.e., location, phone lock-screen, and PPG recordings; for a list of all features, see Tables A1-3 in the appendix). For data that was sampled multiple times per day, we tested different time windows (e.g., from 4 to 48 hours)

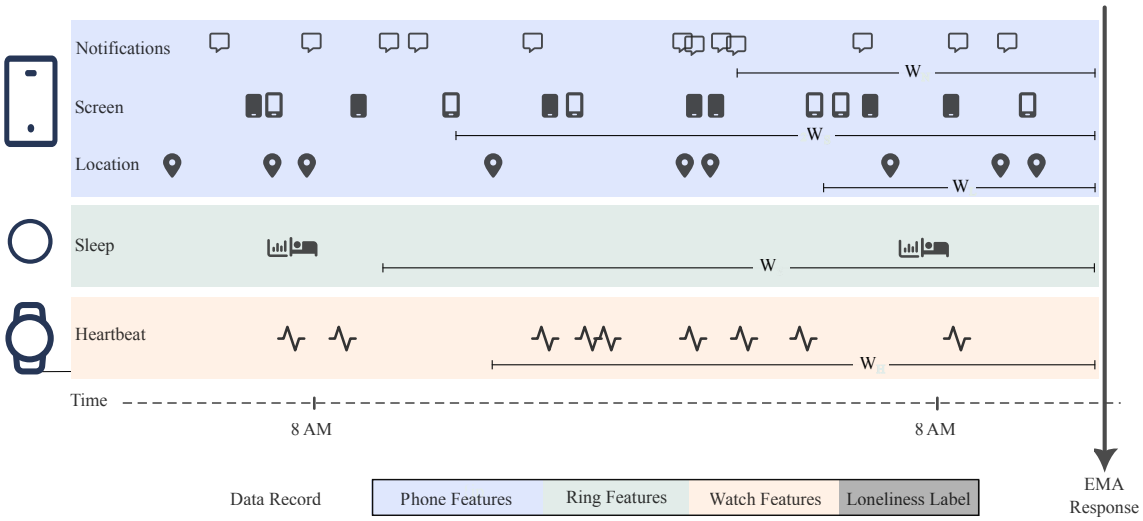

Figure 2: Schematic presentation of feature extraction window and data record compilation from modalities. Each modality provides different sensors and separate window lengths could be defined for extracting features from each of these sensors.

for aggregating the values in order to pair with the subjective responses (i.e., self-reported loneliness used for labeling). Figure 2 shows how different window lengths for each modality have been used to compile data records. After extracting mentioned feature values given a time window, we selected the optimum time window for each feature based on the correlation with the target questionnaire response values.

2.5. *Missing data*

For intensive longitudinal studies assessing experiences in real-time, many factors could interrupt continuous data collection and result in missing data. These factors may result from participants forgetting to wear their devices or charge them overnight, or may result due to technical issues (e.g., server congestion or permission issues over the phone). Using the monitoring tools in the data collection server, we were able to track most of these issues over the course of study and solve them in the shortest time possible. After all these arrangements, having a certain degree of missing data is inevitable and must be taken care of in the analysis stage. In our analysis, we tested two data imputation methods to handle missing data. We replaced the missing values with the average of: A) two preceding and succeeding valid samples, or B) all valid values [49] of that participant. The latter method yielded a higher correlation with the target questionnaire response values for the majority of the features. Therefore, for consistency, we used this method to handle missing data for all of the features.

2.6. *Classification*

We defined the binary classification labels (true/false values) based on the participant's self-reported loneliness rating (scaled from 0-100, 0 = not at all, 100 = extremely) as below or above the median to have a balanced number of labels. We developed a random forest [50] method to predict these classification labels using the 84 selected features. Random Forests is a machine learning algorithm that combines multiple decision trees to make more accurate predictions. Leveraging multiple decision trees enables this algorithm to process high dimensional data efficiently. Before training predictive models, we Z-normalized the features using each participant's data in the training set to reduce interpersonal bias in feature space. For model evaluation, we used a one-subject-out policy for predictive modeling evaluation, where we used the most recent 50% of each participant's data as a test set and the remaining data of that participant with all the other participants for the training model. We calculated the accuracy, F1-score, precision, recall, and mean squared error for all the test entries combined. We defined true positive and false positive predictions if our model detected the loneliness class respectively correct and incorrect (according to EMA labels). Similarly,

true negative and false negative predictions were instances where the model did not detect the loneliness class while they were correct and incorrect, respectively. Given these terms we can define aforementioned performance metrics as:

$$Accuracy = \frac{number\ of\ true\ negative + number\ of\ true\ positive}{total\ number\ of\ data\ records}$$

$$Precision = \frac{number\ of\ true\ positive}{number\ of\ true\ positive\ +\ number\ of\ false\ positive}$$

$$Recall = \frac{number\ of\ true\ positive}{number\ of\ true\ positive\ +\ number\ of\ false\ negative}$$

$$F1 = \frac{2\ *\ precision\ *\ recall}{precision\ +\ recall}$$

$$MeanSquaredError = \frac{\sum_{total\ data\ records} number\ of\ false\ negative\ +\ number\ of\ falsepositives}{total\ number\ f\ data\ records}$$

2.7. Feature Importance

A limitation of using classic model evaluation metrics is that they lack insight into the dynamics of prediction. To address this limitation, the SHapley Additive exPlanations (SHAP) method was used to investigate the contribution of features to predictions given a model [51]. SHAP offers explainability and insight into the contribution of each feature used by a given model in the prediction stage. Given a trained model, this method generates numerical estimates (called SHAP values) that represent how much a feature value affects the output value of the model and what direction (toward either of the classification labels) this effect is. Variations of this method have been proposed for different machine learning models. Some of them consider the trained model as black-box and extract local explanations using efficient sampling in the feature space [52]. Other variations of SHAP analysis methods are proposed for specific machine learning models. In this work, we leveraged path dependent feature perturbation algorithms [53] developed for tree-based models. This method splits the feature space while recursively following the decision path for a given datapoint. This approach enabled us to enhance the accuracy and interpretability of our models, making them more effective tools for decision-making and analysis.

3. Results

3.1. Aim 1: Loneliness detection overall performance

We first present the performance of our proposed loneliness detection method enabled by using the multiple ubiquitous sensing devices AWARE, Oura ring, and Samsung watch. As previously mentioned, we conducted a two-month monitoring study with 30 participants who completed five self-report assessments per day, totaling to about 7,300 data points. The trained model was evaluated by comparing the estimated loneliness values with the corresponding ground truth values (i.e., collected by subjective questionnaires). This procedure was repeated for every participant; in other words, 30 different personal models were built. In the following, we report the aggregated results obtained from all the trained models. Our loneliness detection models obtained an accuracy of 82% Table 2, precision score of 81%, and value of 0.83 Area Under Curve (AUC). Higher AUC generally represents better classification capability (ranging from 0.5 showing random classification to 1 that is for a perfect classification). Figure 3 shows the confusion matrix of the loneliness detection models extracted from about 3,600 tested samples. The obtained true positives (detecting loneliness correctly) and true negatives (estimating no loneliness correctly) were considerably lower compared to the false positives and false negatives.

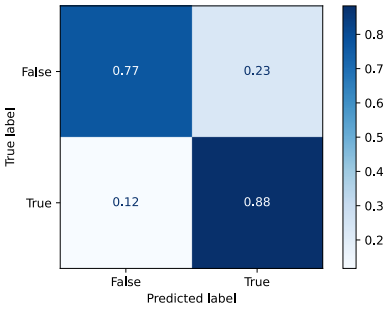

Figure 3: Confusion matrix of loneliness prediction model performance using all channels.

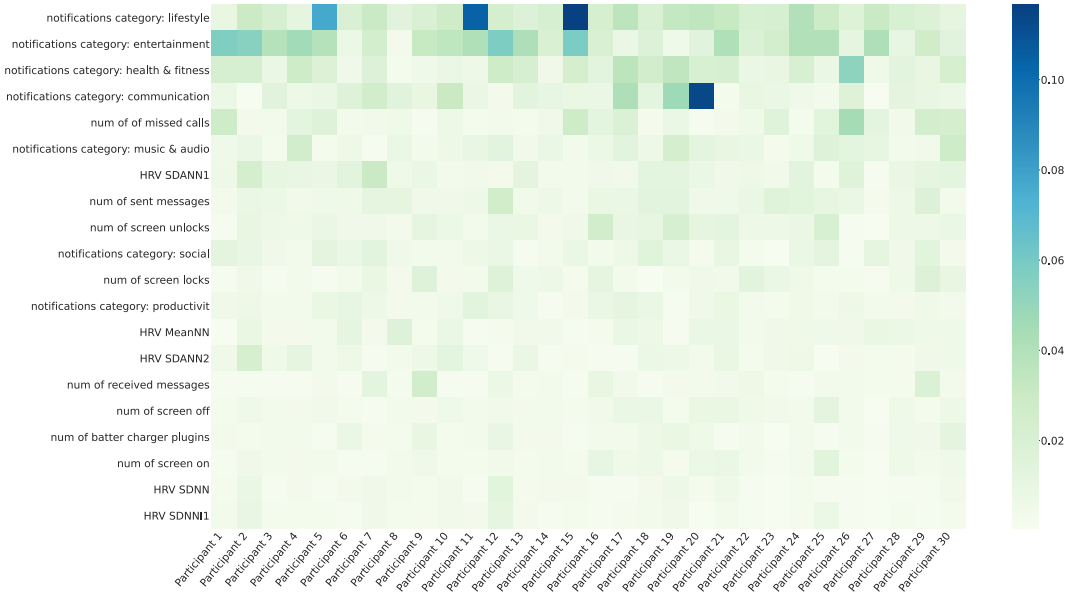

Figure 4: mean of absolute SHAP values for each participant-feature pairs. The top 20 features with the highest average values across participants are presented in a sorted order. $HRVSDANN1$, $HRVSDANN1$ is the standard deviation of average RR intervals extracted from segments of PPG collected by the smart watch for 1-minute and 5-minute intervals respectively. Please refer to Appendix A.4 for a full description of all features.

3.2. Aim 2: Explainability and Feature Importance Analysis

As discussed, we obtained loneliness detection accuracy of 82% using all the modalities from three devices (smartwatch, ring, and phone). In this section, we explore further into the analysis to investigate the effect of the features on the performance of loneliness detection using explainable machine learning techniques. To this end, we carried out SHapley Additive exPlanations [51] (SHAP) analysis to gain deeper insight into the detection models. SHAP values serve as a measure showing the significance of each feature in the model's ability to perform the detection. As previously mentioned, we trained a personal model for each participant. To assess the impact of each feature on the loneliness detection among participants, we calculated the mean SHAP values across the test samples.

Figure 4 depicts the absolute SHAP values of the 20 features for each participant, with dark blue and light green hues showing highest and lowest values, respectively. Overall, the 6 most influential features from our prediction model were related to behavioral markers extracted from smartphones (i.e., the AWARE platform), while the HRV features – collected from the smart watch – had less impact. However, the influence of each feature varied across participants. For example, the number of notifications from lifestyle and communication applications had a significant impact on Participants 15 and 20, respectively, whereas these

features did not exert as similar of an impact on the remaining participants. This inter-individual difference indicates the importance of personalization in loneliness detection methods.

The findings shown in Figure 4 further offer a broad understating of the impact of individual features on the loneliness detection. However, due to averaging the absolute values, this representation obscures the distribution and direction of the effects of each feature on each participant. Specifically, a positive SHAP value denotes that the corresponding feature shows a positive impact on the loneliness class detection, while a negative value contributes to the non-loneliness class detection. Thus, Figure 5a,5b shows the positive and negative SHAP values of two randomly selected participants, revealing the variation in the importance of the features. Loneliness class is represented on the right side of the y axis in this Figure, while non-loneliness is represented on the left side. The red and blue colors illustrate negative associations and positive associations respectively for a given feature. For example, physiological features were the most influential for Participant 8 (Figure 5a), whereas Participant 10's most important features (Figure 5b) were behavioral features. It also shows that on average, Participant 8's lower values of HRV metrics (represented by the blue color) are associated with higher feelings of loneliness (left side of the vertical line). These figures also indicate how the order and impact of features vary between the participants for loneliness detection. For example, for Participant 8, the most influential features are extracted from the smartwatch. However, this effect is less evident for Participant 10 (Figure 5b). For instance, decreases in the HRV CVSD feature (presented in blue color) results in loneliness for Participant 10. This is while higher values of this feature (red) results in non-loneliness class.

3.3. Aim 3: Loneliness Detection using Different Sets of Devices

We investigated the effects of different modalities on the performance of our loneliness detection method. To this end, in the training phase, we built four random forest models, three of which were trained with data collected from a single device (i.e., smart ring, smart watch, or smartphone), and one of which was trained using data from all three devices. Subsequently, the models were evaluated using test data from the same device(s). This assessment strategy allowed us to select the optimal set of devices based on the desired criteria, including performance, costs, or user burden. The performance of the models is indicated in Table 2. The smart ring (Oura) is small, lightweight, and easy-to-use, with a battery life of approximately one week following a full charge. The device provides sleep quality, physical activity, and nocturnal HR and RMSSD parameters with a rather high accuracy [54, 55, 35]. Although Oura obtained the highest recall value compared to all other cases, it showed the poorest accuracy, F1-score, and precision for loneliness detection. The second-best accuracy of a single device was obtained using the smart watch (Samsung), which enabled us to acquire HR and multiple HRV measurements at all times. In comparison to Oura, the watch obtained better precision but worse recall. The AWARE framework, which only uses smartphone logging and sensing features, obtained the highest accuracy and precision but the lowest recall compared to the Oura ring and Samsung watch. AWARE runs as a background application on the user's smartphone. This passive sensing eliminates the need for user input, in contrast to wearable devices that require continuous wear. It should be noted however, that the AWARE data collection is limited to the behavioral and contextual parameters. For instance, our previous study [56] indicated that COVID-19 lockdown had an adverse impact on such passive smartphone-based data collection, as user mobility decreased (e.g., fewer location changes). Our results showed that multimodality has a positive impact on the overall performance of loneliness detection, providing an improved trade-off between precision and recall. Specifically, the three devices used in the study together obtained higher precision values compared to the Oura ring alone, and slightly higher recall values compared to Samsung Watch and AWARE alone.

4. Discussion

Advances in technology offer exciting future direction for mental health prevention and treatment. With the availability of passive sensing and assessments of people's experiences, behaviors, and mental states in real-time and with greater ecological validity, researchers and medical professionals can have a better understanding of individualized experiences of mental health. Loneliness is an important predictor of physical

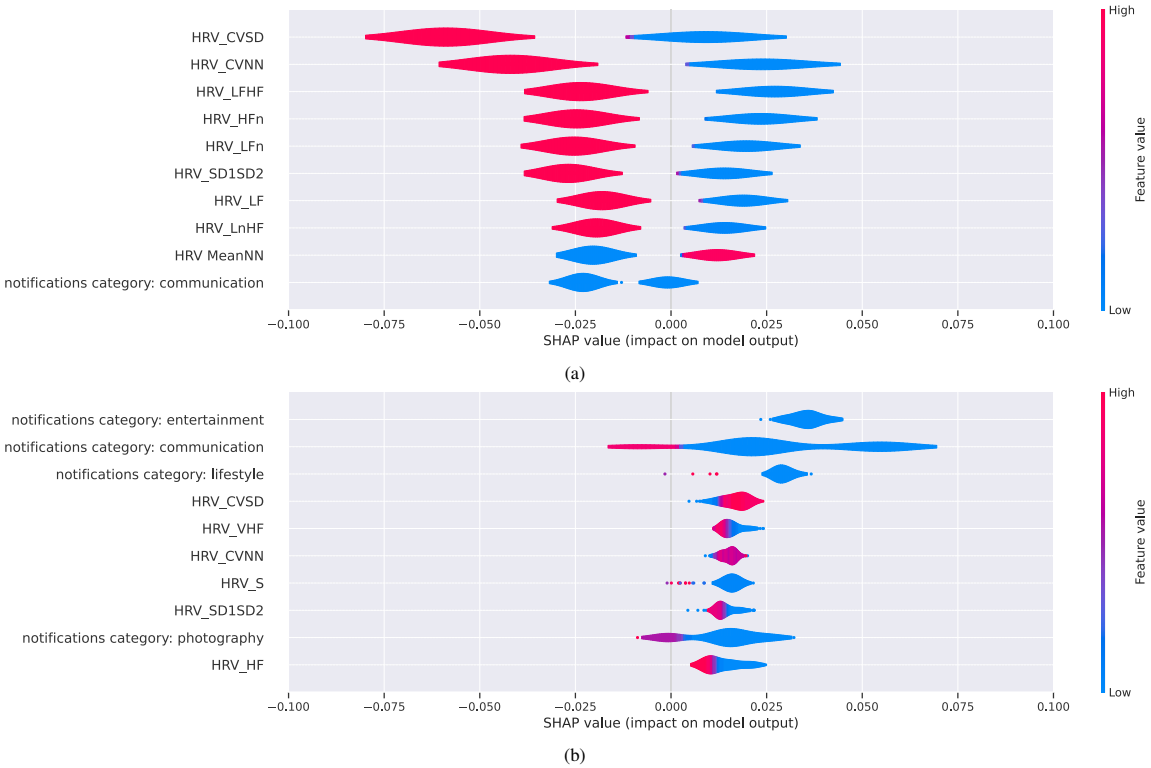

Figure 5: Summary of SHAP-based explainers for top 10 features for Participants 8 (a) 10 (b). HRV data are extracted from watch signals. HRV_CVSD is RMSSD divided by the mean of the RR intervals (MeanNN). HRV_LF and HRV_HF are the spectral power of low frequencies (0.04 to 0.15 Hz) and high frequencies (0.15 to 0.4 Hz) respectively. HRV_LFn and HRV_HFn are power-normalized versions of HRV_LF and HRV_HF features. SDNN the standard deviation of the RR intervals. CVNN is the standard deviation of the RR intervals (SDNN) divided by the mean of the RR intervals. Please refer to Appendix for the full list of features description.

Table 2: The performance of the loneliness detection methods using different devices

Device(s)	Accuracy	Precision	Recall	F1 score	MSE
Oura Ring	0.565	0.556	0.988	0.711	0.434
Samsung Watch	0.781	0.809	0.780	0.794	0.220
AWARE (Smartphone)	0.810	0.877	0.756	0.812	0.189
Samsung Watch + Oura Ring + Aware	0.822	0.883	0.773	0.824	0.177

health and mortality [57]. Furthermore, loneliness is associated with several sleep and physiological factors [58]. The U.S. has previously announced a loneliness epidemic [59] and several recent studies are focusing on loneliness across different populations [60]. To address loneliness requires a multimodal approach. Our study adopted a comprehensive approach using multiple devices to passively and continuously monitor people's experiences in order to predict loneliness and assess how loneliness might look across individuals.

Our findings in this study showed that loneliness detection using multimodal assessment with commercially available devices could yield an accuracy of 82%. We compare our proposed ML-based loneliness detection method with current state-of-the-art methods for fully and objectively (i.e., only using sensors from wearable and mobile devices) identifying loneliness. To the best of our knowledge, there exists only one article in the literature which attempts to objectively detect instantaneous loneliness (i.e., not daily or weekly classification) while having an acceptable sample size ($N \geq 10$) . It should be noted that in order to utilize a predictive model in just-in-time adaptive interventions (JITAI) [61], it is critical to deploy models with the ability of instantaneous (fine-grained) predictions. Wu et. al. [32] conducted a three-week data collection of 129 individuals using smartphones. The EMA was collected randomly up to four times a day in which participants were asked to select 4 options of loneliness levels. Using a random forest classifier, they obtained the average AUC of 0.73-0.74 for loneliness detection, which is 10% lower than our method's AUC. Although their results showed how contextual data collected from smartphone sensors could be used in loneliness detection, they did not leverage physiological measurement and multi-modal assessment in their study. As presented in the following, physiological indicators are rich sources of information for loneliness detection models when coupled with contextual markers.

With respect to feature importance for predicting loneliness, we found that across devices, assessments of people's phone behaviors using the AWARE app had the highest average values with loneliness across all participants. More specifically, people's use of their phones either through notifications or engagement with phone apps and calls were highly associated with loneliness. In this respect, our findings do support the existing literature that has primarily used mobile app assessments of loneliness using people's geolocation and phone interactions with others [33, 27, 32]. HRV parameters collected through the Samsung smartwatch were second in their prediction weights on loneliness.

Surprisingly, sleep measures as assessed through the ring had less weight as predictors of loneliness. This finding is somewhat inconsistent with existing literature. Past studies that have used subjective indices of sleep suggest a link between reported sleep disturbances and loneliness [40, 41]. Furthermore, studies that have used objective indicators of sleep (such as polysomnography, sleep watches, and neural scans) are consistent with subjective indices, where sleep quality and deprivation was also found to be associated with social withdrawal and loneliness [62]. One potential limitation of our work that may help to explain our finding is that our model included several features from assessments collected during the day, whereas the Oura ring was the only device that captured bedtime features. While behavioral and some of the activity data are being generated throughout the day, sleep data has values per day. This difference in the frequency of sleep data could suggest further exploration on specific data fusion techniques.

With respect to the nuances across behavioral features and cardiovascular features, one possible explanation for why behavioral features weighed more heavily is the growing literature examining the importance of context [63, 64, 65]. For example, the links across emotions, physiological assessments, and health vary widely across different racial/ethnic groups [66, 67]. Assessing these objective markers lack meaning if not measured with the context of the individual. Loneliness is highly related to a person's perceived social interactions or lack thereof. If an individual feels slight changes in their interactions with others, whether it be in their use of phone apps or virtual engagement with others, it might have changes in their loneliness. Therefore, in having a more comprehensive approach to understanding how loneliness occurs, it still is needed to ground it in the context of social interactions and relationships.

Our aim was to obtain the best performance for objective loneliness detection using all available devices. In other words, in addition to assessing the performance of loneliness detection methods, it is vital to evaluate how each device performs individually. Through this process, researchers can gain insights on how to improve other metrics, such as feasibility and usability within the context of remote health monitoring studies. This information can assist researchers in selecting the optimal set of devices according to their specific requirements. Our findings indicate that multimodality yields a beneficial effect on the overall performance

of loneliness detection, offering an enhanced balance between precision and recall. Additionally, multimodality offers competitive data fusion [68] in loneliness detection, which can enhance service availability. This enhancement is achieved through the utilization of three independent battery-powered devices, which ensures that loneliness detection remains operational even if one or two of the devices become unavailable due to factors such as battery depletion, technical problems, or decreased user mobility. Moreover, multimodality provides complementary fusion of sensing modalities [68], which can improve the explainability of the analysis. Each device is not directly dependent on other devices, but the data can be aggregated to provide a more complete depiction of loneliness under observation. The Oura ring collects physical activity and nocturnal health parameters, including sleep quality and HR. The Samsung watch provides nighttime and daytime physiological parameters, including HR and HRV. AWARE enables us to acquire behavioral information (e.g., social interactions) and contextual information (e.g., location). In other words, each device assesses different phenomena, and assessing all of these at the same time, in real-time, can provide a more comprehensive view of how loneliness occurs for each individual. By including each device in analyses, a potentially more accurate examination of associations between loneliness and diverse health and well-being parameters can be tested. Therefore, it can provide holistic actionable insights for health providers. Overall, multimodality is a promising approach for improving the performance, availability, and explainability of loneliness detection methods, however, it increases the cost and user burden.

5. Conclusion

In this paper we proposed a method for detecting loneliness in a multimodal fully objective setup using smart phones and commercially available wearable devices. We tested our method in a 2-month long study and showed that we can detect loneliness in a continuous way. Our results showed that smartphones and activity-related information during the day are among the most important features for most of the participants. Leveraging multi-modal assessments and passive sensing to continuously and objectively monitor human experiences can help reduce participant burden and support on-going efforts to design personalized treatment and programming to support well-being. Assessing human experiences in real-time can allow us to take a more preventative approach for health treatment rather than a reactive approach. The work conducted in this study was restricted to a demographic of college students and limited to a 2 month period, which therefore suggests further investigation for confirming the generalizability. We are planning to further investigate device-specific feature extraction techniques and personalized modeling for our proposed method.

References

[1] J. Holt-Lunstad, The potential public health relevance of social isolation and loneliness: Prevalence, epidemiology, and risk factors, Public Policy & Aging Report 27 (4) (2017) 127–130.

[2] D. Perlman, L. A. Peplau, Toward a social psychology of loneliness, Personal relationships 3 (1981) 31–56.

[3] N. Leigh-Hunt, D. Bagguley, K. Bash, V. Turner, S. Turnbull, N. Valtorta, W. Caan, An overview of systematic reviews on the public health consequences of social isolation and loneliness, Public health 152 (2017) 157–171.

[4] J. Holt-Lunstad, T. B. Smith, M. Baker, T. Harris, D. Stephenson, Loneliness and social isolation as risk factors for mortality: a meta-analytic review, Perspectives on psychological science 10 (2) (2015) 227–237.

[5] R. B. Olsen, J. Olsen, F. Gunner-Svensson, B. Waldstrøm, Social networks and longevity. a 14 year follow-up study among elderly in denmark, Social science & medicine 33 (10) (1991) 1189–1195.

[6] D. Banerjee, M. Rai, Social isolation in covid-19: The impact of loneliness (2020).

[7] J. Leland, How loneliness is damaging our health (Apr 2022).
URL `https://www.nytimes.com/2022/04/20/nyregion/loneliness-epidemic.html`

[8] F. Bu, A. Steptoe, D. Fancourt, Who is lonely in lockdown? cross-cohort analyses of predictors of loneliness before and during the covid-19 pandemic, Public health 186 (2020) 31–34.

[9] C. M. Lee, J. M. Cadigan, I. C. Rhew, Increases in loneliness among young adults during the covid-19 pandemic and association with increases in mental health problems, Journal of Adolescent Health 67 (5) (2020) 714–717.

[10] W. D. Killgore, S. A. Cloonan, E. C. Taylor, D. A. Lucas, N. S. Dailey, Loneliness during the first half-year of covid-19 lockdowns, Psychiatry research 294 (2020) 113551.

[11] M. Luchetti, J. H. Lee, D. Aschwanden, A. Sesker, J. E. Strickhouser, A. Terracciano, A. R. Sutin, The trajectory of loneliness in response to covid-19., American Psychologist 75 (7) (2020) 897.

[12] G. Prati, A. D. Mancini, The psychological impact of covid-19 pandemic lockdowns: a review and meta-analysis of longitudinal studies and natural experiments, Psychological medicine 51 (2) (2021) 201–211.

[13] M. Ernst, D. Niederer, A. M. Werner, S. J. Czaja, C. Mikton, A. D. Ong, T. Rosen, E. Brähler, M. E. Beutel, Loneliness before and during the covid-19 pandemic: A systematic review with meta-analysis., American Psychologist 77 (5) (2022) 660.

[14] A. Steptoe, A. Shankar, P. Demakakos, J. Wardle, Social isolation, loneliness, and all-cause mortality in older men and women, Proceedings of the National Academy of Sciences 110 (15) (2013) 5797–5801.

[15] J. Beller, A. Wagner, Disentangling loneliness: differential effects of subjective loneliness, network quality, network size, and living alone on physical, mental, and cognitive health, Journal of Aging and Health 30 (4) (2018) 521–539.

[16] J. T. Cacioppo, S. Cacioppo, Older adults reporting social isolation or loneliness show poorer cognitive function 4 years later, Evidence-based nursing 17 (2) (2014) 59–60.

[17] B. Yu, A. Steptoe, Y. Chen, X. Jia, Social isolation, rather than loneliness, is associated with cognitive decline in older adults: the china health and retirement longitudinal study, Psychological medicine 51 (14) (2021) 2414–2421.

[18] J. Beller, A. Wagner, Loneliness, social isolation, their synergistic interaction, and mortality., Health Psychology 37 (9) (2018) 808.

[19] A. C. Patterson, G. Veenstra, Loneliness and risk of mortality: A longitudinal investigation in alameda county, california, Social science & medicine 71 (1) (2010) 181–186.

[20] C. Park, A. Majeed, H. Gill, J. Tamura, R. C. Ho, R. B. Mansur, F. Nasri, Y. Lee, J. D. Rosenblat, E. Wong, et al., The effect of loneliness on distinct health outcomes: a comprehensive review and meta-analysis, Psychiatry Research 294 (2020) 113514.

[21] L. A. Rico-Uribe, F. F. Caballero, N. Martín-María, M. Cabello, J. L. Ayuso-Mateos, M. Miret, Association of loneliness with all-cause mortality: A meta-analysis, PloS one 13 (1) (2018) e0190033.

[22] L. J. Levine, M. A. Safer, Sources of bias in memory for emotions, Current directions in psychological science 11 (5) (2002) 169–173.

[23] A. A. Stone, J. E. Schwartz, J. M. Neale, S. Shiffman, C. A. Marco, M. Hickcox, J. Paty, L. S. Porter, L. J. Cruise, A comparison of coping assessed by ecological momentary assessment and retrospective recall., Journal of personality and social psychology 74 (6) (1998) 1670.

[24] D. L. Thomas, E. Diener, Memory accuracy in the recall of emotions., Journal of personality and social psychology 59 (2) (1990) 291.

[25] M. van Winkel, M. Wichers, D. Collip, N. Jacobs, C. Derom, E. Thiery, I. Myin-Germeys, F. Peeters, Unraveling the role of loneliness in depression: the relationship between daily life experience and behavior, Psychiatry 80 (2) (2017) 104–117.

[26] L. C. Hawkley, M. H. Burleson, G. G. Berntson, J. T. Cacioppo, Loneliness in everyday life: cardiovascular activity, psychosocial context, and health behaviors., Journal of personality and social psychology 85 (1) (2003) 105.

[27] K. J. MacDonald, G. Willemsen, D. I. Boomsma, J. A. Schermer, Predicting loneliness from where and what people do, Social Sciences 9 (4) (2020) 51.

[28] A. S. Bello-Valle, A. Martínez-Rebollar, W. Sánchez, H. Estrada-Esquivel, A predictive model for automatic detection of loneliness and social isolation using machine learning, Computación y Sistemas 26 (1) (2022) 113–124.

[29] M. S. Ahmed, N. Ahmed, Less is more: Leveraging digital behavioral markers for real-time identification of loneliness in resource-limited settings.

[30] M. M. Qirtas, E. Zafeiridi, D. Pesch, E. B. White, Loneliness and social isolation detection using passive sensing techniques: Scoping review, JMIR mHealth and uHealth 10 (4) (2022) e34638.

[31] D. Prabhu, M. Kholghi, M. Sandhu, W. Lu, K. Packer, L. Higgins, D. Silvera-Tawil, Sensor-based assessment of social isolation and loneliness in older adults: A survey, Sensors 22 (24) (2022) 9944.

[32] C. Wu, A. N. Barczyk, R. C. Craddock, G. M. Harari, E. Thomaz, J. D. Shumake, C. G. Beevers, S. D. Gosling, D. M. Schnyer,

Improving prediction of real-time loneliness and companionship type using geosocial features of personal smartphone data, Smart Health 20 (2021) 100180.
[33] A. Doryab, D. K. Villalba, P. Chikersal, J. M. Dutcher, M. Tumminia, X. Liu, S. Cohen, K. Creswell, J. Mankoff, J. 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) (2019) e13209.
[34] Y. Gao, A. Li, T. Zhu, X. Liu, X. Liu, How smartphone usage correlates with social anxiety and loneliness, PeerJ 4 (2016) e2197.
[35] R. Cao, I. Azimi, F. Sarhaddi, H. Niela-Vilen, A. Axelin, P. Liljeberg, A. M. Rahmani, Accuracy assessment of oura ring nocturnal heart rate and heart rate variability in comparison with electrocardiography in time and frequency domains: comprehensive analysis, Journal of Medical Internet Research 24 (1) (2022) e27487.
[36] M. Nissen, S. Slim, K. Jäger, M. Flaucher, H. Huebner, N. Danzberger, P. A. Fasching, M. W. Beckmann, S. Gradl, B. M. Eskofier, et al., Heart rate measurement accuracy of fitbit charge 4 and samsung galaxy watch active2: Device evaluation study, JMIR Formative Research 6 (3) (2022) e33635.
[37] N. C. Jacobson, B. Summers, S. Wilhelm, Digital biomarkers of social anxiety severity: digital phenotyping using passive smartphone sensors, Journal of medical Internet research 22 (5) (2020) e16875.
[38] D. Ben-Zeev, E. A. Scherer, R. Wang, H. Xie, A. T. Campbell, Next-generation psychiatric assessment: Using smartphone sensors to monitor behavior and mental health., Psychiatric rehabilitation journal 38 (3) (2015) 218.
[39] R. Wang, F. Chen, Z. Chen, T. Li, G. Harari, S. Tignor, X. Zhou, D. Ben-Zeev, A. 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, 2014, pp. 3–14.
[40] S. C. Griffin, A. B. Williams, S. G. Ravyts, S. N. Mladen, B. D. Rybarczyk, Loneliness and sleep: A systematic review and meta-analysis, Health psychology open 7 (1) (2020) 2055102920913235.
[41] M. A. Hom, C. Chu, M. L. Rogers, T. E. Joiner, A meta-analysis of the relationship between sleep problems and loneliness, Clinical Psychological Science 8 (5) (2020) 799–824.
[42] I. Moura, A. Teles, D. Viana, J. Marques, L. Coutinho, F. Silva, Digital phenotyping of mental health using multimodal sensing of multiple situations of interest: A systematic literature review, Journal of Biomedical Informatics (2022) 104278.
[43] G. Vashisht, R. Vashisht, A study on the tizen operating system, International Journal of Computer Trends and Technology 12 (1) (2014) 14–15.
[44] D. Ferreira, V. Kostakos, A. K. Dey, Aware: mobile context instrumentation framework, Frontiers in ICT 2 (2015) 6.
[45] [link].
URL `https://play.google.com/store/apps/details?id=org.healthscitech.thrive`
[46] M. Feli, I. Azimi, A. Anzanpour, A. M. Rahmani, P. Liljeberg, An energy-efficient semi-supervised approach for on-device photoplethysmogram signal quality assessment, Smart Health 28 (2023) 100390.
[47] Y. Wang, I. Azimi, K. Kazemi, A. M. Rahmani, P. Liljeberg, Ppg signal reconstruction using deep convolutional generative adversarial network, in: 2022 44th Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC), IEEE, 2022, pp. 3387–3391.
[48] K. Kazemi, J. Laitala, I. Azimi, P. Liljeberg, A. M. Rahmani, Robust ppg peak detection using dilated convolutional neural networks, Sensors 22 (16) (2022) 6054.
[49] J. M. Engels, P. Diehr, Imputation of missing longitudinal data: a comparison of methods, Journal of clinical epidemiology 56 (10) (2003) 968–976.
[50] L. Breiman, Random forests, Machine learning 45 (2001) 5–32.
[51] S. M. Lundberg, S.-I. Lee, A unified approach to interpreting model predictions, Advances in neural information processing systems 30.
[52] E. Strumbelj, I. Kononenko, An efficient explanation of individual classifications using game theory, The Journal of Machine Learning Research 11 (2010) 1–18.
[53] S. M. Lundberg, G. Erion, H. Chen, A. DeGrave, J. M. Prutkin, B. Nair, R. Katz, J. Himmelfarb, N. Bansal, S.-I. Lee, From local explanations to global understanding with explainable ai for trees, Nature machine intelligence 2 (1) (2020) 56–67.
[54] M. A. Mehrabadi, I. Azimi, F. Sarhaddi, A. Axelin, H. Niela-Vilén, S. Myllyntausta, S. Stenholm, N. Dutt, P. Liljeberg, A. M. Rahmani, et al., Sleep tracking of a commercially available smart ring and smartwatch against medical-grade actigraphy in everyday settings: instrument validation study, JMIR mHealth and uHealth 8 (11) (2020) e20465.
[55] H. Niela-Vilen, I. Azimi, K. Suorsa, F. Sarhaddi, S. Stenholm, P. Liljeberg, A. M. Rahmani, A. Axelin, Comparison of oura smart ring against actigraph accelerometer for measurement of physical activity and sedentary time in a free-living context, CIN: Computers, Informatics, Nursing.
[56] S. Jafarlou, J. Lai, I. Azimi, Z. Mousavi, S. Labbaf, R. C. Jain, N. Dutt, J. L. Borelli, A. Rahmani, et al., Objective prediction of next-day's affect using multimodal physiological and behavioral data: Algorithm development and validation study, JMIR Formative Research 7 (1) (2023) e39425.
[57] T. J. Holwerda, T. G. Van Tilburg, D. J. Deeg, N. Schutter, R. Van, J. Dekker, M. L. Stek, A. T. Beekman, R. A. Schoevers, Impact of loneliness and depression on mortality: results from the longitudinal ageing study amsterdam, The British Journal of Psychiatry 209 (2) (2016) 127–134.
[58] T. Matthews, A. Danese, A. M. Gregory, A. Caspi, T. E. Moffitt, L. Arseneault, Sleeping with one eye open: loneliness and sleep quality in young adults, Psychological medicine 47 (12) (2017) 2177–2186.
[59] D. V. Jeste, E. E. Lee, S. Cacioppo, Battling the modern behavioral epidemic of loneliness: suggestions for research and interventions, JAMA psychiatry 77 (6) (2020) 553–554.
[60] D. L. Surkalim, M. Luo, R. Eres, K. Gebel, J. van Buskirk, A. Bauman, D. Ding, The prevalence of loneliness across 113 countries: systematic review and meta-analysis, bmj 376.
[61] D. Spruijt-Metz, W. Nilsen, Dynamic models of behavior for just-in-time adaptive interventions, IEEE Pervasive Computing

13 (3) (2014) 13–17.
[62] E. Ben Simon, M. P. Walker, Sleep loss causes social withdrawal and loneliness, Nature communications 9 (1) (2018) 3146.
[63] K. H. Greenaway, E. K. Kalokerinos, L. A. Williams, Context is everything (in emotion research), Social and Personality Psychology Compass 12 (6) (2018) e12393.
[64] I. Kim, H. Goh, N. Narziev, Y. Noh, U. Lee, Understanding user contexts and coping strategies for context-aware phone distraction management system design, Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies 4 (4) (2020) 1–33.
[65] A. Mehrotra, S. R. Müller, G. M. Harari, S. D. Gosling, C. Mascolo, M. Musolesi, P. J. Rentfrow, Understanding the role of places and activities on mobile phone interaction and usage patterns, Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies 1 (3) (2017) 1–22.
[66] B. Mesquita, M. Boiger, J. De Leersnyder, Doing emotions: The role of culture in everyday emotions, European Review of Social Psychology 28 (1) (2017) 95–133.
[67] D. R. Williams, S. A. Mohammed, Discrimination and racial disparities in health: evidence and needed research, Journal of behavioral medicine 32 (2009) 20–47.
[68] W. Elmenreich, An introduction to sensor fusion, Vienna University of Technology, Austria 502 (2002) 1–28.

Appendix A. Full List of Used Features

Table A.3: List of features extracted from smart ring

Category	Feature
Activity	Average activity level during the whole day.
	Number of kilocalories one spent in a day based on the amount of physical activity
	The total number of kilocalories one consumed in a day
	Every 5 minutes a character is shown to indicate the level of intensity of an activity.
	How many minutes of high intensity activity was performed in a day.
	Number of minutes where one is not doing anything in a day.
	Number of minutes where one is doing low intensity activity in a day.
	Number of minutes where one is doing medium intensity activity in a day.
	The average amount of energy it costs to complete a task for each minute of the span of one activity period.
	Total number of minutes of when the user was not wearing the ring during the day
	Total number of minutes spent sleeping during a day.
	Number of steps taken during the day
Sleep Features	the amount of time spent awake during your sleep period
	average beats per minute
	percentage of sleep period spent asleep
	heart rate at 5 minutes each sleep period
	average heart rate during the sleep period
	lowest heart rate during sleep period (resting heart rate)
	The time it takes a person to fall asleep after turning the lights out
	Percentage of sleep time when the user was moving
	Measure of the beat-to-beat differences in heart rate
	Measures the beat-to-beat differences in heart rate in the first five minutes of the sleep period
	The measure of the change in your body temperature while you sleep
	The difference between your body's core temperature and the temperature of your sleeping environment
	Represents the amount of time a person spends asleep during a given period of time

Table A.4: List of features extracted from smart watch

ID	Feature
HRV_MeanNN	The mean of the RR intervals.
HRV_SDNN	The standard deviation of the RR intervals.
HRV_SDANN1	The standard deviation of average RR intervals extracted from 1-minute segments of time series data
HRV_SDNNI1	The mean of the standard deviations of RR intervals extracted from 1-minute segments of time series data
HRV_RMSSD	The square root of the mean of the squared successive differences between adjacent RR intervals.
HRV_SDSD	The standard deviation of the successive differences between RR intervals.
HRV_CVNN	The standard deviation of the RR intervals (SDNN) divided by the mean of the RR intervals (MeanNN).
HRV_CVSD	The root mean square of successive differences (RMSSD) divided by the mean of the RR intervals (MeanNN).
HRV_MedianNN	The median of the RR intervals.
HRV_MadNN	The median absolute deviation of the RR intervals.
HRV_MCVNN	The median absolute deviation of the RR intervals (MadNN) divided by the median of the RR intervals (MedianNN).
HRV_IQRNN	The interquartile range (IQR) of the RR intervals.
HRV_PNN50	The proportion of RR intervals greater than 50ms, out of the total number of RR intervals.
HRV_PNN20	The proportion of RR intervals greater than 20ms, out of the total number of RR intervals.
HRV_HTI	The HRV triangular index, measuring the total number of RR intervals divided by the height of the RR intervals histogram.
HRV_TINN	The baseline width of the RR intervals distribution obtained by triangular interpolation, where the error of least squares determines the triangle. It is an approximation of the RR interval distribution.
HRV_LF	The spectral power of low frequencies (0.04 to 0.15 Hz).
HRV_HF	The spectral power of high frequencies (0.15 to 0.4 Hz).
HRV_VHF	The spectral power of very high frequencies (0.4 to .5 Hz).
HRV_LFHF	The ratio obtained by dividing the low frequency power by the high frequency power.
HRV_LFn	The normalized low frequency, obtained by dividing the low frequency power by the total power.
HRV_HFn	The normalized high frequency, obtained by dividing the low frequency power by the total power.
HRV_LnHF	The log transformed HF.
HRV_SD1	Standard deviation perpendicular to the line of identity. It is an index of short-term RR interval fluctuation
HRV_SD2	Standard deviation along the identity line. Index of long-term HRV changes.
HRV_SD1SD2	Ratio of SD1 to SD2. Describes the ratio of short term to long term variations in HRV.
HRV_S	Area of ellipse described by SD1 and SD2 (pi * SD1 * SD2)
HR	Heart rate

Table A.5: List of features extracted from smartphone

Category	**Feature**
Call	Outgoing call duration
	Income call counts
	Outgoing call counts
	Missedcall counts
	Voicemail counts
Notification	Number of notifications from applications type: Productivity
	Number of notifications from applications type: Photography
	Number of notifications from applications type: Communication
	Number of notifications from applications type: Lifestyle
	Number of notifications from applications type: Social
	Number of notifications from applications type: Shopping
	Number of notifications from applications type: Health & Fitness
	Number of notifications from applications type: Entertainment
	Number of notifications from applications type: Music & Audio
Messages	Number of received messages
	Number of sent messagess
Screen	Number of screen off
	Number of screen on
	Number of screen locks
	Number of screen unlocks
Location (GPS)	Variance of latitude
	Variance of speed
	Mean of speed
	Number of places
	Home duration
	Outdoor duration
	Mean of outdoor duration
	Standard deviation of outdoor duration
	Type of the place with longest duration other than home
	Total travel distance
Battery	Number of battery charger plugins